

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
11 March 2004 (11.03.2004)

PCT

(10) International Publication Number
WO 2004/019893 A2

- (51) International Patent Classification⁷: **A61K**
- (21) International Application Number:
PCT/US2003/027523
- (22) International Filing Date:
2 September 2003 (02.09.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
10/231,956 30 August 2002 (30.08.2002) US
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 10/231,956 (CIP)
Filed on 30 August 2002 (30.08.2002)
- (71) Applicant (for all designated States except US): **RIGEL PHARMACEUTICALS, INC.** [US/US]; 1180 Veterans Blvd., South San Francisco, CA 94080 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **LORENS, James, B.**

[US/NO]; Kraakenesveien 10, N-5152 Bones (NO). **BO-GENBERGER, Jakob** [DE/US]; 3028 Central Avenue, Ventura, CA 93003 (US). **HOLLAND, Sacha** [GB/US]; 518 Foerster Street, Apt. B, San Francisco, CA 94127 (US). **XU, Weiduan** [CN/US]; 1800 22nd Avenue, Apt. 203, San Francisco, CA 94122 (US).

(74) Agents: **KELLY, Beth, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, 8th Floor, San Francisco, CA 94111 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,

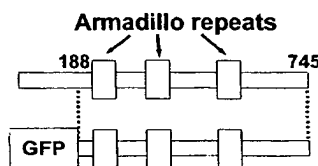
[Continued on next page]

(54) Title: MODULATORS OF ANGIOGENESIS

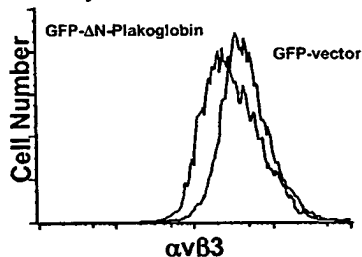
GFP-ΔN-Plakoglobin Expression in Endothelial Cells Downregulates αvβ3 Surface Levels

Functional Screening Hit

Plakoglobin
GFP-ΔN-Plakoglobin



FACS Analysis



(57) Abstract: The present invention relates to regulation of angiogenesis. More particularly, the present invention is directed to nucleic acids encoding "angiogenesis regulatory proteins and nucleic acids" which are involved in modulation of angiogenesis. The invention further relates to methods for identifying and using agents, including small organic molecules, antibodies, peptides, cyclic peptides, nucleic acids, antisense nucleic acids, RNAi, and ribozymes, that modulate angiogenesis via modulation of angiogenesis regulatory proteins and nucleic acids; as well as to the use of expression profiles and compositions in diagnosis and therapy of diseases related to angiogenesis.

BEST AVAILABLE COPY

WO 2004/019893 A2



SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations* AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU,

TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*
- *of inventorship (Rule 4.17(iv)) for US only*

Published:

- *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Modulators of Angiogenesis

CROSS-REFERENCES TO RELATED APPLICATIONS

This application is a continuation-in-part of Application No. 10/231,956 filed
5 August 30, 2002, which is herein incorporated by reference for all purposes.

FIELD OF THE INVENTION

The present invention relates to regulation of angiogenesis. More particularly,
the present invention is directed to nucleic acids encoding "angiogenesis regulatory proteins
10 and nucleic acids" which are involved in modulation of angiogenesis. The invention further
relates to methods for identifying and using agents, including small organic molecules,
antibodies, peptides, cyclic peptides, nucleic acids, antisense nucleic acids, RNAi, and
ribozymes, that modulate angiogenesis via modulation of angiogenesis regulatory proteins
and nucleic acids; as well as to the use of expression profiles and compositions in diagnosis
15 and therapy of diseases related to angiogenesis.

BACKGROUND OF THE INVENTION

Angiogenesis is typically limited in a normal adult to the placenta, ovary,
endometrium and sites of wound healing. However, angiogenesis, or its absence, plays an
20 important role in the maintenance of a variety of pathological states. Some of these states are
characterized by neovascularization, *e.g.*, cancer, diabetic retinopathy, glaucoma, and age
related macular degeneration. More specifically, pathological states characterized by
neovascularization include lymphoma, hematologic cancers, melanoma, breast cancer, lung
cancer, prostate cancer, colon cancer, ovarian cancer, liver cancer, Kaposi's sarcoma,
25 metastatic disease, rheumatoid arthritis, psoriasis, and benign proliferative disorders
including hemangiomas. Others, *e.g.*, stroke, infertility, heart disease, ulcers, delayed wound
healing, and scleroderma, are diseases of angiogenic insufficiency. Therefore, there is a need
to identify nucleic acids encoding proteins involved in the regulation of angiogenesis, to
identify, *e.g.*, modulators of angiogenesis, as well as new therapeutic and diagnostic
30 applications.

BRIEF SUMMARY OF THE INVENTION

The present application identifies, for the first time, a number of proteins and DNA molecules involved in regulation of angiogenesis, *e.g.*, angiogenesis regulatory proteins and DNA molecules. The invention further relates to methods for identifying and using agents, including small organic molecules, antibodies, peptides, cyclic peptides, nucleic acids, antisense nucleic acids, RNAi, and ribozymes, that modulate angiogenesis via modulation of angiogenesis regulatory proteins and DNA molecules; as well as to the use of expression profiles and compositions in diagnosis and therapy of diseases related to insufficient or increased angiogenesis.

In one aspect, the present invention provides a method for identifying a compound that modulates angiogenesis, the method comprising the steps of:

(i) contacting the compound with an angiogenesis regulatory nucleic acid, or an angiogenesis regulatory polypeptide or a fragment thereof encoded by a nucleic acid, wherein the nucleic acid hybridizes under stringent conditions to a second nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523; and (ii) determining the functional effect of the compound upon the nucleic acid or polypeptide.

In one embodiment, the functional effect is determined in vitro. In another embodiment, the functional effect is a physical effect. In another embodiment, the functional effect is determined by measuring ligand or substrate binding to the polypeptide. In another embodiment, the functional effect is a chemical effect. In another embodiment, the functional effect is determined by measuring an enzymatic activity.

In one embodiment, the polypeptide is expressed in a eukaryotic host cell. In a further embodiment, the functional effect is a physical effect. In another embodiment, the functional effect is determined by measuring ligand or substrate binding to the polypeptide. In a further embodiment, the functional effect is a chemical or phenotypic effect. In another embodiment, the functional effect is determined by measuring an enzymatic activity. In another embodiment, the host cell is an endothelial cell. In a further embodiment, the functional effect is determined by measuring $\alpha v \beta 3$ expression or haptotaxis, or chemotaxis, or co-culture tube formation.

In one embodiment, modulation is inhibition of angiogenesis.

In one embodiment, the polypeptide is recombinant. In another embodiment, the nucleic acid comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523.

In another embodiment, the polypeptide comprises a sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID

NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:520, SEQ ID NO:522, and SEQ ID NO:524.

In one embodiment, the compound is an antibody, an antisense molecule, a small organic molecule, or a peptide.

5 In another aspect, the present invention provides a method for identifying a compound that modulates angiogenesis, the method comprising the steps of (i) contacting the compound with a nucleic acid, or a polypeptide or a fragment thereof encoded by a nucleic acid, wherein the nucleic acid hybridizes under stringent conditions to a second nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:3, SEQ
10 ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID
15 NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID
20 NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523; (ii) determining the functional effect of the compound upon the nucleic acid or polypeptide; and (iii) expressing the nucleic acid or polypeptide in a cell, contacting the nucleic acid or polypeptide with the compound, and determining the
25 phenotypic or chemical effect upon the cell.

In another aspect, the present invention provides a method of modulating angiogenesis in a subject, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified as a modulator of angiogenesis using the methods described herein. In one embodiment, the subject is a human. In a further
30 embodiment, the compound is an antibody, an antisense molecule, a small organic molecule, a peptide, or an RNAi molecule. In another embodiment, the compound inhibits angiogenesis.

In another aspect, the present invention provides a method of modulating angiogenesis in a subject, the method comprising the step of administering to the subject a

therapeutically effective amount of a polypeptide, the polypeptide encoded by a nucleic acid that hybridizes under stringent conditions to a second nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:63, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:165, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:218, SEQ ID NO:232, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:340, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, and SEQ ID NO:498.

In another aspect, the present invention provides a method of modulating angiogenesis in a subject, the method comprising the step of administering to the subject a therapeutically effective amount of a nucleic acid that hybridizes under stringent conditions to a second nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523.

In another aspect, the present invention provides the nucleic acid sequences of SEQ ID NO:165, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, and SEQ ID NO:330.

In another aspect, the present invention provides the amino acid sequences of SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, and SEQ ID NO:331.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides a schematic diagram of an assay for modulators of angiogenesis.

Figure 2 provides results of an experiment demonstrating the effect of a plakoglobin-GFP fusion protein expression on levels of the cell surface marker $\alpha\beta 3$.

Plakoglobin is an exemplar angiogenesis regulatory protein. Human umbilical vein endothelial (HUVEC) cells were transfected with a vector expressing the plakoglobin-GFP fusion protein or a control vector, expressing GFP only. Cells were incubated with APC-labeled antibodies directed against the cell surface marker $\alpha\beta 3$. The X-axis depicts cell number and the Y-axis depicts the amount of $\alpha\beta 3$ -APC antibody derived fluorescence.

Cells transfected with the plakoglobin-GFP fusion protein construct exhibit lower $\alpha\beta 3$ expression levels than control cells.

Figure 3 also provides results of an experiment demonstrating the effect of a plakoglobin-GFP fusion protein (*i.e.*, GFP- Δ N-plakoglobin) expression on levels of the cell surface marker $\alpha\beta 3$. The figure also depicts the presence of armadillo repeats in the functional screen hit, GFP- Δ N-plakoglobin.

Figure 4 demonstrates that GFP- Δ N-Plako expression downregulates vitronectin receptors ($\alpha\beta 3$ and $\alpha\beta 5$) but not the fibronectin receptor ($\alpha 5\beta 1$).

Figure 5 demonstrates that expression of GFP- Δ N-Plakoglobin downregulates endogenous catenin levels, as well as expression of endogenous plakoglobin and VE-cadherin.

Figure 6 demonstrates that GFP- Δ N-Plakoglobin expression upregulates surface levels of the laminin receptor, $\alpha 6\beta 4$.

Figure 7 provides provides results of an experiment demonstrating the effect of a HoxB2-GFP fusion protein expression (*i.e.*, Δ N-HoxB2) on levels of the cell surface marker $\alpha\beta 3$. Cells transfected with the Δ N-HoxB2-GFP fusion protein construct exhibit lower $\alpha\beta 3$ expression levels than control cells. The figure also provides the structure of the functional screen hit, demonstrating that the homeodomain is present in Δ N-HoxB2.

Figure 8 demonstrates that Δ N-HoxB2 downregulates both α and $\beta 3$ integrin subunit surface expression, as well as $\alpha\beta 5$ surface expression.

Figure 9 demonstrates that Δ N-HoxB2 Inhibits Proliferation in HUVECs, a primary cell line.

Figure 10 demonstrates that full length HoxB2 phenocopies Δ N-HoxB2 in the downregulation of $\alpha\beta 3$.

5 Figure 11 provides provides results of an experiment demonstrating the effect of a SUSP-1 GFP fusion protein expression (*i.e.*, GFP-SUSP-1) on levels of the cell surface marker $\alpha\beta 3$. Cells transfected with the GFP-SUSP-1 fusion protein construct exhibit lower $\alpha\beta 3$ expression levels than control cells. The figure also provides the structure of the functional screen hit, demonstrating that GFP-SUSP-1 is an N-terminal fusion in the sense
10 orientation and that a cDNA insert is present before the wild-type start codon.

Figure 12 demonstrates that the GFP-SUSP-1 screening hit does not affect proliferation when expressed in HUVEC (human umbilical vein endothelial cells), PASMC (pulmonary artery smooth muscle cells), and NHDF (normal human dermal fibroblasts).

15 Figure 13 provides expression analysis of SUSP-1 mRNA in a variety of cell lines.

Figure 14 demonstrates that SUSP-1 siRNAs inhibit SUSP-1 mRNA expression.

Figure 15 provides a new sequence determination for a nucleic acid encoding an ABC transporter protein.

20 Figure 16 provides a new sequence determination for a nucleic acid encoding an HSPA5 protein.

Figure 17 provides a new sequence determination for a nucleic acid encoding a protein disulfide isomerase.

25 Figure 18 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 1.

Figure 19 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 3.

Figure 20 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 17

30 Figure 21 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 8.

Figure 22 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 9.

Figure 23 provides a new sequence determination for a nucleic acid encoding a 1/226 protein.

Figure 24 provides a new sequence determination for a nucleic acid encoding an FLJ10688 protein.

5 Figure 25 provides a new sequence determination for a nucleic acid encoding a KIAA1583 protein.

Figure 26 provides a new sequence determination for a nucleic acid encoding a KIAA1814 protein.

10 Figure 27 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 4.

Figure 28 provides a new sequence determination for a nucleic acid encoding a peroxidasin/melanoma antigen related protein.

Figure 29 provides a new sequence determination for a nucleic acid encoding a WD40/SOCS box protein.

15 Figure 30 provides provides results of an experiment demonstrating the effect of a $\Delta 5FADS$ antisense screening hit (*i.e.*, GFP- $\Delta 5FADS$ or D5ADS) on levels of the cell surface marker $\alpha v \beta 3$. Cells transfected with the GFP- $\Delta 5FADS$ antisense construct exhibit lower $\alpha v \beta 3$ expression levels than control cells. The figure also provides the structure of the functional screen hit, demonstrating that GFP- $\Delta 5FADS$ is in the antisense orientation and
20 that a the anti sense portion of the $\Delta 5FADS$ nucleic acid is about 300 base pairs long.

Figure 31 demonstrates that the GFP-D5FADS antisense screening hit does not affect cell proliferation.

Figure 32 demonstrates that $\Delta 5FADS$ siRNAs inhibit $\Delta 5FADS$ mRNA expression.

25 Figure 33 demonstrates that $\Delta 5FADS$ siRNAs inhibit haptotaxis.

Figure 34 demonstrates that $\Delta 5FADS$ siRNAs inhibit SDF-1 induced chemotaxis in Jurkat cells.

Figure 35 demonstrates that $\Delta 6FADS$ is a functional screening hit. The figure shows that $\Delta 6FADS$ transcript is spliced to the KIAA1583 functional screen hit.

30 Figure 36 demonstrates tha $\Delta 6FADS$ mRNA levels are straongly reduced in cells that express the antisense KIAA1583 functional screen hit.

Figure 37 that $\Delta 6FADS$ siRNAs inhibit $\Delta 6FADS$ mRNA expression and also reduce $\alpha v \beta 3$ surface expression.

DETAILED DESCRIPTION OF THE INVENTION

Introduction

Using a functional assay, angiogenesis regulatory proteins and nucleic acids have been identified and cloned from a green fluorescent protein (GFP) fusion library.

5 Primary endothelial cells were transduced with retroviral cDNA/GFP fusion libraries. Transduced cells were selected and assayed for alterations consistent with inhibition of angiogenesis, e.g. downregulation of cell surface expression of $\alpha v \beta 3$ or downregulation of haptotaxis.

10 Using the screen described above, fifty-two proteins and/or DNA molecules were shown for the first time to regulate angiogenesis. Thirty-four of the proteins and/or DNA molecules are known, while fifteen of the proteins and/or DNA molecules are completely novel. However, this is the first demonstration of a function in regulation of angiogenesis for these molecules.

15 The ABC transporter is involved in peroxisome biogenesis. Relevant sequence data for the protein, nucleic acids encoding the ABC transporter, and related sequences include the nucleic acid accession number NM_000033, SEQ ID NO:3; protein accession number NP_000024, SEQ ID NO:4; a new sequence determination, SEQ ID NO:2; and the functional screen hit, SEQ ID NO:1. The protein has an ATP-binding domain. The gene encoding the ABC transporter is linked to a genetic disease, adrenoleukodystrophy
20 (Mosser *et al.*, *Nature* 361:682-3 (1993)). An antisense version of the ABC transporter gene, e.g., SEQ ID NO:1, was identified as a functional hit in the screen for modulators of angiogenesis.

25 PCB1, a poly (rC) binding protein functions as a translational co-activator and also binds 3' UTRs of RNA, most likely through KH domains, thereby regulating RNA stability. Relevant sequence data for the protein, nucleic acids encoding PCB1, and related sequences include the nucleic acid accession number NM_006196, SEQ ID NO:32; protein accession number NP_006187, SEQ ID NO:33; and the functional screen hit, SEQ ID NO:5-31. An antisense version of the gene encoding PCB1 was identified as a functional hit in the screen for modulators of angiogenesis.

30 SLC1A5, an amino acid transporter, has ten transmembrane spanning segments. Relevant sequence data for the protein, nucleic acids encoding PCB1, and related sequences include the nucleic acid accession number NM_005628, SEQ ID NO:43; protein accession number NP_005619, SEQ ID NO:44; and the functional screen hit, SEQ ID

NO:34-42. An antisense version of the gene encoding SLC1A5 was identified as a functional hit in the screen for modulators of angiogenesis.

Chromobox homolog 6 is involved in chromatin regulation. Relevant sequence data for the protein, nucleic acids encoding chromobox homolog 6, and related sequences include the nucleic acid accession number NM_014292, SEQ ID NO:57; protein accession number NP_055107, SEQ ID NO:58; and the functional screen hit, SEQ ID NO:44-56. An antisense version of chromobox homolog 6 was identified as a functional hit in the screen for modulators of angiogenesis.

Cytochrome C oxidase subunit 1 is encoded by a mitochondrial gene and transfers electrons from cytochrome C to O₂ (Ingman *et al.*, *Nature* 408:708-713 (2000); (Leninger, *Principles of Biochemistry* (1984); Stryer, *Biochemistry* (1995)) Relevant sequence data for the protein, nucleic acids encoding cytochrome C oxidase, and related sequences include the nucleic acid accession number NC_001807, SEQ ID NO:63; protein accession number NP_536845, SEQ ID NO:64; and the functional screen hit, SEQ ID NO:59-62. A sense version of the gene encoding cytochrome C oxidase, encoding a 21 amino acid peptide, was identified as a functional hit in the screen for modulators of angiogenesis.

Δ -5 fatty acid desaturase (FADSD5 or Δ 5FADS) introduces double bonds into a fatty acyl chain. The protein is involved in arachadonic acid synthesis and has a fatty acid desaturase domain and a heme-binding domain. Relevant sequence data for the protein, nucleic acids encoding Δ -5 fatty acid desaturase, and related sequences include the nucleic acid accession number NM_013402, SEQ ID NO:68; protein accession number NP_037534, SEQ ID NO:69; and the functional screen hit, SEQ ID NO:65-67. An antisense version of Δ -5 fatty acid desaturase gene was identified as a functional hit in the screen for modulators of angiogenesis.

The dynactin p27 subunit protein was shown to be an angiogenesis regulatory protein. The protein has an RGD domain. Relevant sequence data for the protein, and the nucleic acid encoding dynactin p27 subunit protein, include the nucleic acid accession number NC_006571, SEQ ID NO:70; and protein accession number NP_006562.1, SEQ ID NO:71. An antisense version of the gene encoding dynactin p27 subunit protein was identified as a functional hit in the screen for modulators of angiogenesis.

Elongation factor 1 alpha (EF1 α) was shown to be involved in regulation of angiogenesis. Relevant sequence data for the protein, nucleic acids encoding EF1 α , and related sequences include the nucleic acid accession number NC_001402, SEQ ID NO:76;

protein accession number NP_001393, SEQ ID NO:77; and the functional screen hit, SEQ ID NO:72-75.

Elongation factor 4 gamma (eIF4 γ) is a translation elongation factor. The protein also has a guanine nucleotide binding domain and a W2 domain. Relevant sequence data for the protein, nucleic acids encoding cytochrome C oxidase, and related sequences include the nucleic acid accession number NM_001418, SEQ ID NO:81; protein accession number NP_001409, SEQ ID NO:82; and the functional screen hit, SEQ ID NO:78-80. An antisense version of the eIF4 γ gene was identified as a functional hit in the screen for modulators of angiogenesis. The antisense hit encodes a 21 amino acid peptide that may also be useful as a therapeutic or to modulate angiogenesis.

The filamin A protein binds integrin and regulates actin polymerization. Relevant sequence data for the protein, nucleic acids encoding filamin A, and related sequences include the nucleic acid accession number NM_001456, SEQ ID NO:86; protein accession number NP_001447, SEQ ID NO:87; and the functional screen hit, SEQ ID NO:83-85. A sense version of the filamin A gene was identified as a functional hit in the screen for modulators of angiogenesis.

Inactivation of the gene encoding the G α 13 protein causes a severe vascular phenotype in mice due to an inhibition of angiogenesis (vasculogenesis was apparent). This protein couples to the thrombin receptor. Thrombin and G α 13 protein activate Rho and tyrosine-phosphorylation of focal adhesion proteins FAK, Paxillin, and p130Cas. G α 13 also interacts with the cytoplasmic domain of adherins, leading to increased levels of b-catenin and transcriptional activation. Relevant sequence data for the protein, the nucleic acid encoding the G α 13 protein, and related sequences include the nucleic acid accession number NM_006572, SEQ ID NO:89; protein accession number NP_006563.1, SEQ ID NO:90; and the functional screen hit, SEQ ID NO:88. A sense version of the G α 13 gene was identified as a functional hit in the screen for modulators of angiogenesis.

The HOXB2 protein is a transcription factor. Relevant sequence data for the protein, nucleic acids encoding HOXB2, and related sequences include the nucleic acid accession number NM_002145, SEQ ID NO:120; protein accession number NP_002136, SEQ ID NO:121; and the functional screen hit, SEQ ID NO:91-119. A sense version of the HOXB2 gene was identified as a functional hit in the screen for modulators of angiogenesis.

HSPA5, a member of the HSP70 family of proteins has a role in protein folding and assembly in the endoplasmic reticulum. Expression of the protein is glucose regulated. The protein has an ATPase domain which is fused in frame to GFP in the

functional screen hit. Relevant sequence data for the protein, nucleic acids encoding HSPA5, and related sequences include the nucleic acid accession number NM_005347, SEQ ID NO:128; protein accession number NP_005338, SEQ ID NO:129; a new sequence determination, SEQ ID NO:127; and the functional screen hit, SEQ ID NO:122-126.

5 Interferon gamma receptor 1 (IFN γ 1) is a cytokine receptor that induces antiangiogenic genes. Relevant sequence data for the protein, nucleic acids encoding IFN γ 1, and related sequences include the nucleic acid accession number NM_000416, SEQ ID NO:139; protein accession number NP_000407, SEQ ID NO:140; and the functional screen hit, SEQ ID NO:130-138. An antisense version of the IFN γ 1 gene was identified as a
10 functional hit in the screen for modulators of angiogenesis.

Importin α 4 (karyopherin α 3) is a subunit of the nuclear localization signal receptor. Relevant sequence data for the protein, nucleic acids encoding importin α 4, and related sequences include the nucleic acid accession number NM_002267, SEQ ID NO:153; protein accession number NP_002258.1, SEQ ID NO:154; and the functional screen hit, SEQ
15 ID NO:141-152. An antisense version of the importin α 4 gene was identified as a functional hit in the screen for modulators of angiogenesis.

The lysosomal pepstatin-insensitive protease (CLN2) is encoded by a gene associated with juvenile neuronal ceroid lipofuscinosis disease. Relevant sequence data for the protein, nucleic acids encoding CLN2, and related sequences include the nucleic acid
20 accession number NM_000391, SEQ ID NO:163; protein accession number NP_000382, SEQ ID NO:164; and the functional screen hit, SEQ ID NO:155-162. An antisense version of the CLN2 gene was identified as a functional hit in the screen for modulators of angiogenesis.

A novel protein disulfide isomerase was shown to have a role in angiogenesis.
25 Relevant sequence data for the nucleic acid encoding the novel protein disulfide isomerase, includes SEQ ID NO:165. An antisense version of gene encoding the novel protein disulfide isomerase was identified as a functional hit in the screen for modulators of angiogenesis. The antisense hit encodes a 45 or 32 amino acid peptide that may also be useful as a therapeutic or to modulate angiogenesis.

30 Microtubule-associated protein 4 (MAP4) binds tubulin and may link microtubules to other elements of the cytoskeleton. See *e.g.*, Chapin & Bulinski, *J. Cell. Sci* 98:27-36 (1991); and Chapin *et al.*, *Biochemistry* 34:2289-2301 (1995). Four isoforms of MAP4 are known. Relevant sequence data for the proteins, nucleic acids encoding MAP4 isoforms, and related sequences include the nucleic acid accession numbers NM_002375,

NM_030884, NM_030885, and NM_030883; SEQ ID NOs:169, 171, 173, and 175; protein accession numbers NP_002366, NP_112146, NP_112147, and NP_112245; SEQ ID NOs:170, 172, 174, and 176; and the functional screen hit, SEQ ID NO:166-168. An antisense version of MAP4 was identified as a functional hit in the screen for modulators of angiogenesis.

Neural precursor cell expressed developmentally down regulated 5 (Nedd5) is a septin GTPase. Relevant sequence data for the protein, nucleic acids encoding Nedd5, and related sequences include the nucleic acid accession number NM_004404, SEQ ID NO:183; protein accession number NM_004395, SEQ ID NO:184; and the functional screen hit, SEQ ID NO:177-182. A sense version of the Nedd5 gene was identified as a functional hit in the screen for modulators of angiogenesis.

A novel tyrosine/serine/threonine kinase was found to be involved in angiogenesis. The protein is also a breast cancer antigen. Relevant sequence data for the protein, nucleic acids encoding the novel kinase, and related sequences include the nucleic acid accession number NM_044379, SEQ ID NO:202; protein accession number XP_044379, SEQ ID NO:203; and the functional screen hit, SEQ ID NO:185-201. A sense version of the gene encoding the novel kinase was identified as a functional hit in the screen for modulators of angiogenesis.

A novel protein encoded by a gene on human chromosome 1 was found to be involved in angiogenesis. The chromosome 1 DNA sequence identified was previously identified as intron 1 of the JAK1 gene. Relevant sequence data for chromosome 1 DNA sequence include a new sequence determination, SEQ ID NO:210; and the functional screen hit, SEQ ID NO:204-209. The functional screen hit is believed to be in the antisense orientation.

A novel protein encoded by a gene on human chromosome 3 was found to be involved in angiogenesis. The chromosome 3 DNA sequence was previously identified as H41 exon 2/3 and part of intron 1. Relevant sequence data for the chromosome 3 DNA sequence include a new sequence determination, SEQ ID NO:218; and the functional screen hit, SEQ ID NO:211-217. The functional screen hit is in the sense orientation.

A novel protein encoded by a gene on human chromosome 17 was found to be involved in angiogenesis. Relevant sequence data for the chromosome 17 DNA sequence include a new sequence determination, SEQ ID NO:227; and the functional screen hit, SEQ ID NO:219-226. The functional screen hit is believed to be in the antisense orientation.

A novel protein encoded by a gene on human chromosome 8 was found to be involved in angiogenesis. Relevant sequence data for the chromosome 8 hit include a new sequence determination, SEQ ID NO:232; and the functional screen hit, SEQ ID NO:228-231. The functional screen hit is believed to be in the antisense orientation or encodes a nine amino acid peptide that may also be useful as a therapeutic or to modulate angiogenesis.

A novel protein encoded by a gene on human chromosome 9 was found to be involved in angiogenesis. The chromosome 9 DNA sequence was previously identified as intronic DNA from the RXRa gene. Relevant sequence data for chromosome 9 hit include a new sequence determination, SEQ ID NO:248; and the functional screen hit, SEQ ID NO:233-247. The functional screen hit is believed to be in the antisense orientation.

A novel protein (1/226) that binds to PAK2 in a yeast two hybrid screen was found to be involved in angiogenesis. Relevant sequence data for the novel 1/226 hit include a new sequence determination, SEQ ID NO:274; and the functional screen hit, SEQ ID NO:249-273. The functional screen hit is in the sense orientation.

Another novel protein FLJ10688 contains fibronectin domains and binds to STAT-6 in a yeast two hybrid screen. Relevant sequence data for the protein, nucleic acids encoding FLJ10688, and related sequences include the nucleic acid accession number NM_018179, SEQ ID NO:286; protein accession number NP_060649, SEQ ID NO:287; a new sequence determination, SEQ ID NO:285; and the functional screen hit, SEQ ID NO:275-284. The functional screen hit is in the sense orientation.

Another novel protein KIAA0217 contains an RNA recognition motif. Relevant sequence data for the protein, nucleic acids encoding KIAA0217, and related sequences include the nucleic acid accession number XM_040265, SEQ ID NO:297; protein accession number NP_040265, SEQ ID NO:298; and the functional screen hit, SEQ ID NO:288-296. The functional screen hit is in the antisense orientation.

Another novel protein KIAA1583 is homologous to PDZ-RGS, which regulates Ephrin B signaling. Relevant sequence data for the protein, nucleic acids encoding KIAA1583, and related sequences include the nucleic acid accession number AB046803, SEQ ID NO:308; protein accession number BAB13409, SEQ ID NO:309; a new sequence determination, SEQ ID NO:307; and the functional screen hit, SEQ ID NO:299-306. The functional screen hit is in the antisense orientation. The KIAA1583 transcript is spliced to the transcript of the $\Delta 6$ -fatty acid desaturase enzyme, also located on chromosome 11q12. The antisense KIAA1583 functional screen hit, inhibits expression of $\Delta 6$ -fatty acid desaturase mRNA. In addition, a $\Delta 6$ -fatty acid desaturase specific siRNA decreased $\Delta 6$ -fatty

acid desaturase message levels and inhibited tube formation in a co-culture assay. Thus, $\Delta 6$ -fatty acid desaturase is also an angiogenesis regulatory protein and the protein and its related nucleic acids can be used in the screening methods of the present invention. Compounds identified in such screens can be used in the therapeutic methods described herein.

5 Representative $\Delta 6$ -fatty acid desaturase amino acid sequence data is found at accession numbers NP_004256, SEQ ID NO:522; and AAD31282, SEQ ID NO:524. Related nucleic acid sequence data is found at accession numbers NM_004265, SEQ ID NO 521; and AF134404, SEQ ID NO:523.

A novel protein, KIAA1814, was found to be involved in angiogenesis.

10 Relevant sequence data for functional screen hit include the nucleic acid accession number XM_046822, SEQ ID NO:318; protein accession number XP_046822, SEQ ID NO:319; a new sequence determination, SEQ ID NO:317; and the functional screen hit, SEQ ID NO:310-316. The functional screen hit is in the sense orientation.

A novel protein encoded by a gene on human chromosome 4 was found to be
15 involved in angiogenesis. Relevant sequence data for functional screen hit include a new sequence determination, SEQ ID NO:320.

Another novel protein identified as an angiogenesis regulatory protein is related to a peroxidasin/melanoma antigen. The protein has peroxidase, vWF, Ig, and leucine rich domains. Relevant sequence data for the protein, nucleic acids encoding the novel
20 protein, and related sequences include the nucleic acid accession number XM_056455, SEQ ID NO:324; protein accession number XP_056455, SEQ ID NO:325; a new sequence determination, SEQ ID NO:323; and the functional screen hit, SEQ ID NO:321-322. The functional screen hit is in the sense orientation comprising a fusion to the peroxidase domain.

Another novel protein identified herein as an angiogenesis regulatory protein
25 is a novel WD40/SOCS box protein. Relevant sequence data for the protein, nucleic acids encoding the novel WD40/SOCS box protein, and related sequences include the nucleic acid accession number NM_018639, SEQ ID NO:330; protein accession number NP_061109, SEQ ID NO:331; a new sequence determination, SEQ ID NO:329; and the functional screen hit, SEQ ID NO:326-328. The functional screen hit is in the antisense orientation.

30 A Numb-like protein contains adapter and PTB domains and is involved in regulation of angiogenesis. Relevant sequence data for the protein, nucleic acids encoding the Numb-like protein, and related sequences include the nucleic acid accession number NM_004756, SEQ ID NO:340; protein accession number NP_004747, SEQ ID NO:341; and

the functional screen hit, SEQ ID NO:332-339. The functional screen hit is in the sense orientation.

Protein p130CAS (Crk-associated substrate) was identified as an angiogenic regulatory protein. p130 CAS is an adapter protein and has SH3 domains. (Dorssers *et al.*, *Mol. Endocrinol.* 7:870-878 (1993); Brinkman *et al.*, *J. Natl. Cancer Inst.* 92:112-120 (2000)). Relevant sequence data for the protein, nucleic acids encoding FLJ10688, and related sequences include the nucleic acid accession number NM_014567, SEQ ID NO:351; protein accession number NP_055382, SEQ ID NO:352; and the functional screen hit, SEQ ID NO:342-350. The orientation of this functional screen hit is antisense.

Perlecan, a basement membrane component, was identified as an angiogenesis regulatory protein. (Nicole *et al.*, *Nat. Genet.* 26:(4) 480-483 (2000)). Relevant sequence data for the protein, nucleic acids encoding perlecan, and related sequences include the nucleic acid accession number NM_005529, SEQ ID NO:365; protein accession number NP_005520, SEQ ID NO:366; and the functional screen hit, SEQ ID NO:353-364. The functional screen hit is in the antisense orientation.

The T cell activation protein PGR-1, was identified as an angiogenesis regulatory protein. Relevant sequence data for the protein, nucleic acids encoding PGR-1, and related sequences include the nucleic acid accession number NM_033296, SEQ ID NO:377; protein accession number NP_150638, SEQ ID NO:378; and the functional screen hit, SEQ ID NO:367-376. The functional screen hit is in the sense orientation.

Phosphoribosyl pyrophosphate synthetase associated protein 1 (PRPSAP1), was identified as an angiogenesis regulatory protein. PRPSAP1 regulates a nucleotide synthesis pathway. Relevant sequence data for the protein, nucleic acids encoding PRPSAP1, and related sequences include the nucleic acid accession number NM_002766, SEQ ID NO:384; protein accession number NP_002757, SEQ ID NO:385; and the functional screen hit, SEQ ID NO:379-383. The functional screen hit is in the sense orientation.

Plakoglobin, a member of the catenin family, was identified as an angiogenesis regulatory protein. Plakoglobin binds to E-cadherin, N-cadherin, and desmoglein proteins. Relevant sequence data for the protein, nucleic acids encoding plakoglobin, and related sequences include the nucleic acid accession numbers NM_002230, SEQ ID NO:406, NM_021991, SEQ ID NO:408; protein accession numbers NP_002221, SEQ ID NO:407, NP_068831, SEQ ID NO:409; and the functional screen hit, SEQ ID NO:386-405. Two functional screen hits were identified: one in the sense orientation, fused to GFP; and one in the antisense orientation.

Plasminogen activator inhibitor 1 (SERPINE1 or PAI-1) was identified as an angiogenesis regulatory protein. PAI-1 is a protease inhibitor. Relevant sequence data for the protein, nucleic acids encoding PAI-1, and related sequences include the nucleic acid accession number NM_000602, SEQ ID NO:419; protein accession number NP_000593, SEQ ID NO:420; and the functional screen hit, SEQ ID NO:410-418. The functional screen hit is in the antisense orientation.

Proteosomal subunit Y, an $\text{INF}\gamma$ regulated proteosomal subunit, was identified as an angiogenesis regulatory protein. Relevant sequence data for proteosomal subunit Y includes the functional screen hit, SEQ ID NO:421. The functional screen hit is in the antisense orientation.

Rap2B was identified as an angiogenesis regulatory protein. Rap2B is a member of the RAS family of GTPases. Relevant sequence data for the protein, nucleic acids encoding Rap2B, and related sequences include the nucleic acid accession number XM_003032, SEQ ID NO:428; protein accession number XP_003032, SEQ ID NO:429; and the functional screen hit, SEQ ID NO:422-427. The functional screen hit is in the antisense orientation.

Semaphorin 3F was identified as an angiogenesis regulatory protein. Semaphorin 3F is a ligand for NP2. Relevant sequence data for the protein, nucleic acids encoding semaphorin 3F, and related sequences include the nucleic acid accession number U38276, SEQ ID NO:437; protein accession number AAB18276, SEQ ID NO:438; and the functional screen hit, SEQ ID NO:430-436. The functional screen hit is in the antisense orientation.

SPARC was identified as an angiogenesis regulatory protein. Relevant sequence data for the protein, nucleic acids encoding SPARC, and related sequences include the nucleic acid accession number NM_003118, SEQ ID NO:439; and protein accession number NP_003109, SEQ ID NO:440.

ssDNA binding protein-1 was identified as an angiogenesis regulatory protein. Relevant sequence data for the protein, nucleic acids encoding ssDNA binding protein-1, and related sequences include the nucleic acid accession number NM_018070, SEQ ID NO:445; protein accession number NP_060540, SEQ ID NO:446; and the functional screen hit, SEQ ID NO:441-444. The functional screen hit is in the sense orientation.

Sumo protease (SUSP-1) was identified as an angiogenesis regulatory protein. Relevant sequence data for the protein, nucleic acids encoding SUSP-1, and related sequences include the nucleic acid accession number NM_015571, SEQ ID NO:456; protein

accession number NP_056386, SEQ ID NO:457; and the functional screen hit, SEQ ID NO:447-455. The functional screen hit is in the sense orientation and includes the 5' untranslated region of the gene.

5 Pantophysin, a synaptophysin homolog, was identified as an angiogenesis regulatory protein (Zhong *et al.*, *Biochim. Biophys. Acta* 1129:235-238 (1992)). Pantophysin was previously shown to be involved in vesicle trafficking. Relevant sequence data for the protein, nucleic acids encoding pantophysin, and related sequences include the nucleic acid accession number NM_006754, SEQ ID NO:462; protein accession number NP_006745, SEQ ID NO:463; and the functional screen hit, SEQ ID NO:458-461. The functional screen
10 hit is in the antisense orientation.

Thrombospondin, a known regulator of angiogenesis, was also identified as a sense hit fused to the laminin domain. Relevant sequence data for the protein, nucleic acids encoding thrombospondin, and related sequences include the nucleic acid accession number NM_0003246, SEQ ID NO:481; protein accession number NP_003237, SEQ ID NO:482; and the functional screen hit, SEQ ID NO:464-480.
15

Transgelin 2 was identified as an angiogenesis regulatory protein. Relevant sequence data for the protein, nucleic acids encoding transgelin 2, and related sequences include the nucleic acid accession number NM_003564, SEQ ID NO:484; protein accession number NP_003555, SEQ ID NO:485; and the functional screen hit, SEQ ID NO:483. The
20 functional screen hit is in the sense orientation.

Vigilin was identified as an angiogenesis regulatory protein. Relevant sequence data for the protein, nucleic acids encoding vigilin, and related sequences include the nucleic acid accession number NM_005336, SEQ ID NO:493; protein accession number NP_005327, SEQ ID NO:494; and the functional screen hit, SEQ ID NO:486-492. The
25 functional screen hit is in the sense orientation.

Vimentin was identified as an angiogenesis regulatory protein. Relevant sequence data for the protein, nucleic acids encoding vimentin, and related sequences include the nucleic acid accession number NM_003380, SEQ ID NO:496, XM_167414, SEQ ID NO:498; protein accession number NP_003371, SEQ ID NO:497, XP_167414, SEQ ID
30 NO:499; and the functional screen hit, SEQ ID NO:495.

Vinexin (SCAM-1) was identified as an angiogenesis regulatory protein. SCAM-1 is an adaptor protein and has an SH3 domain. Relevant sequence data for the protein, nucleic acids encoding SCAM-1, and related sequences include the nucleic acid accession number NM_005775, SEQ ID NO:519; protein accession number NP_005766,

SEQ ID NO:520; and the functional screen hit, SEQ ID NO:500-518. The functional screen hit is in the antisense orientation.

Some of the functional screen hits can be grouped together on the basis of other cellular functions. For example, plakoglobin, the Gα13 protein, and p130CAS protein all appear to function in cell adhesion. The Δ-5 fatty acid desaturase, the ALD protein (ABC transporter), cytochrome C oxidase, and the novel peroxidase protein all may be important in arachadonic acid synthesis. The novel protein disulfide isomerase may be involved in collagen biosynthesis. Pantophysin, HSPA5, and dynactin p27, may all be involved in vesicle trafficking. Without wishing to be bound by theory, regulation of any of these cellular functions may influence the angiogenesis pathway. In addition, without wishing to be bound by theory, both the proteins and/or the nucleic acids identified in the assays are regulators of angiogenesis, *e.g.*, by antisense activity, by overexpression of a functional protein, or as dominant negative inhibitors, or small peptide inhibitors.

Angiogenesis assays described herein reveal for the first time that expression of a nucleic acid molecule encoding the above described angiogenesis regulatory proteins exerted a negative effect on αvβ3 surface expression. In addition, endothelial cells expressing the sequences were strongly inhibited in their haptotactic response to vitronectin, which is an indicator of an anti-angiogenic phenotype. Without wishing to be bound by theory, it appears that the angiogenesis regulatory proteins or fragments of those proteins, peptides derived from the proteins, or peptides and inhibitory DNA or RNA molecules derived from DNA encoding the proteins, provide an anti-angiogenic phenotype. Thus, in addition to their use in screens for modulators of angiogenesis, the angiogenesis regulatory proteins or fragments of those proteins, peptides derived from the proteins, or peptides and inhibitory DNA or RNA molecules derived from DNA encoding the proteins, can also be used as therapeutics for treatment of neovascularization or for treatment of angiogenic insufficiency.

The angiogenesis regulatory proteins, nucleic acids, and other members of the angiogenesis pathway therefore represent targets for the development of angiogenic drugs, preferably anti-angiogenic drugs, *e.g.*, anti-angiogenic drugs for treatment of neovascularization. Neovascularization can be caused by pathologic states including cancer, diabetic retinopathy, glaucoma, age related macular degeneration, lymphoma, hematologic cancers, melanoma, breast cancer, lung cancer, prostate cancer, colon cancer, ovarian cancer, liver cancer, Kaposi's sarcoma, metastatic disease, rheumatoid arthritis, psoriasis, and benign proliferative disorders including hemangiomas. The angiogenesis regulatory proteins,

nucleic acids, and other members of the angiogenesis pathway also represent targets for the development of angiogenic drugs for treatment of angiogenic insufficiency, e.g., stroke, infertility, heart disease, ulcers, delayed wound healing, and scleroderma, are diseases of angiogenic insufficiency. Modulators include small organic molecules, nucleic acids, peptides, cyclic peptides, antibodies, antisense molecules, RNAi molecules, and ribozymes. The nucleic acids and proteins of the invention are also useful for diagnostic applications, using, e.g., nucleic acid probes, oligonucleotides, and antibodies.

Definitions

By “disorder associated with angiogenesis” or “disease associated with angiogenesis” herein is meant a disease state which is marked by either an excess or a deficit of vessel development. Angiogenesis disorders associated with increased angiogenesis include, but are not limited to, cancer and proliferative diabetic retinopathy. Pathological states for which it may be desirable to increase angiogenesis include stroke, heart disease, infertility, ulcers, and scleredema. An increase in angiogenesis may also be desirable in transplantation or for artificial or *in vitro* growth of organs.

The following table (Table 1) presents the angiogenesis regulatory proteins and nucleic acids identified in a screen for such proteins and nucleic acids. Accession numbers and SEQ ID NOs are also provided for each angiogenesis regulatory protein. While the angiogenesis regulatory proteins and nucleic acids are referred to as SEQ ID NOs throughout the present application, the term SEQ ID NO encompasses orthologues and homologues of each angiogenesis regulatory proteins or nucleic acids identified by a SEQ ID NO.

TABLE 1: ANGIOGENESIS REGULATORY PROTEINS AND NUCLEIC ACIDS

Name	Nucleic Acid Accession Number	Nucleic Acid SEQ ID NO	Protein Accession Number	Protein SEQ ID NO
ABC transporter	NM_000033	SEQ ID NO:3	NP_000024	SEQ ID NO:4
PCB1	NM_006196	SEQ ID NO:32	NP_006187	SEQ ID NO:33
SLC1A5	NM_005628	SEQ ID NO:43	NP_005619	SEQ ID NO:44
Chromobox homolog 6	NM_014292	SEQ ID NO:57	NP_055107	SEQ ID NO:58
Cytochrome C	NC_001807	SEQ ID NO:63	NP_536845	SEQ ID NO:64

Name	Nucleic Acid Accession Number	Nucleic Acid SEQ ID NO	Protein Accession Number	Protein SEQ ID NO
oxidase subunit 1				
Δ -5 fatty acid desaturase	NM_013402	SEQ ID NO:68	NP_037534	SEQ ID NO:69
dynactin p27 subunit protein	NC_006571	SEQ ID NO:70	NP_006562.1	SEQ ID NO:71
Elongation factor 1 alpha	NC_001402	SEQ ID NO:76	NP_001393	SEQ ID NO:77
Elongation factor 4 gamma	NM_001418	SEQ ID NO:81	NP_001409	SEQ ID NO:82
filamin A protein	NM_001456	SEQ ID NO:86	NP_001447	SEQ ID NO:87
Gal3 protein	NM_006572	SEQ ID NO:89	NP_006563.1	SEQ ID NO:90
HOXB2 protein	NM_002145	SEQ ID NO:120	NP_002136	SEQ ID NO:121
HSPA5	NM_005347	SEQ ID NO:128	NP_005338	SEQ ID NO:129
Interferon gamma receptor 1	NM_000416	SEQ ID NO:139	NP_000407	SEQ ID NO:140
Importin α 4	NM_002267	SEQ ID NO:153	NP_002258.1	SEQ ID NO:154
lysosomal pepstatin- insensitive protease (CLN2)	NM_000391	SEQ ID NO:163	NP_000382	SEQ ID NO:164
novel protein disulfide isomerase		SEQ ID NO:165		
Microtubule- associated protein 4 (MAP4)	NM_002375	SEQ ID NO:169	NP_002366	SEQ ID NO:170
	NM_030884	SEQ ID NO:171	NP_112146	SEQ ID NO:172
	NM_030885	SEQ ID NO:173	NP_112147	SEQ ID NO:174
	NM_030883	SEQ ID NO:175	NP_112245	SEQ ID NO:176

Name	Nucleic Acid Accession Number	Nucleic Acid SEQ ID NO	Protein Accession Number	Protein SEQ ID NO
Nedd5	NM_004404	SEQ ID NO:183	NM_004395	SEQ ID NO:184
novel tyr/ser/thr kinase	NM_044379	SEQ ID NO:202	XP_044379	SEQ ID NO:203
human chromosome 1 protein		SEQ ID NO:210		
human chromosome 3 protein		SEQ ID NO:218		
human chromosome 17 protein		SEQ ID NO:227		
human chromosome 8 protein		SEQ ID NO:232		
human chromosome 9 protein		SEQ ID NO:248		
1/226		SEQ ID NO:274		
FLJ10688	NM_018179	SEQ ID NO:286	NP_060649	SEQ ID NO:287
KIAA0217	XM_040265	SEQ ID NO:297	NP_040265	SEQ ID NO:298
KIAA1583	AB046803	SEQ ID NO:308	BAB13409	SEQ ID NO:309
KIAA1814	XM_046822	SEQ ID NO:318	XP_046822	SEQ ID NO:319
human chromosome 4 protein		SEQ ID NO:320		
peroxidasin- melanoma antigen related protein	XM_056455	SEQ ID NO:324	XP_056455	SEQ ID NO:325

Name	Nucleic Acid Accession Number	Nucleic Acid SEQ ID NO	Protein Accession Number	Protein SEQ ID NO
WD40/SOCS box protein	NM_018639	SEQ ID NO:330	NP_061109	SEQ ID NO:331
Numb-like protein	NM_004756	SEQ ID NO:340	NP_004747	SEQ ID NO:341
p130CAS	NM_014567	SEQ ID NO:351	NP_055382	SEQ ID NO:352
Perlecan	NM_005529	SEQ ID NO:365	NP_005520	SEQ ID NO:366
PGR-1	NM_033296	SEQ ID NO:377	NP_150638	SEQ ID NO:378
PRPSAP1	NM_002766	SEQ ID NO:384	NP_002757	SEQ ID NO:385
Plakoglobin	NM_002230, NM_021991	SEQ ID NO:406 SEQ ID NO:408	NP_002221 NP_068831	SEQ ID NO:407 SEQ ID NO:409
SERPINE1 or PAI-1	NM_000602	SEQ ID NO:419	NP_000593	SEQ ID NO:420
Proteosomal subunit Y		SEQ ID NO:421		
Rap2B	XM_003032	SEQ ID NO:428	XP_003032	SEQ ID NO:429
Semaphorin 3F	U38276	SEQ ID NO:437	AAB18276	SEQ ID NO:438
SPARC	NM_003118	SEQ ID NO:439	NP_003109	SEQ ID NO:440
ssDNA binding protein-1	NM_018070	SEQ ID NO:445	NP_060540	SEQ ID NO:446
SUSP-1	NM_015571	SEQ ID NO:456	NP_056386	SEQ ID NO:457
Pantophysin	NM_0003246	SEQ ID NO:481	NP_003237	SEQ ID NO:482
Transgelin 2	NM_003564	SEQ ID NO:484	NP_003555	SEQ ID NO:485
Vigilin	NM_005336	SEQ ID NO:493	NP_005327	SEQ ID NO:494
Vimentin	NM_003380 XM_167414	SEQ ID NO:496 SEQ ID NO:498	NP_003371 XP_167414	SEQ ID NO:497 SEQ ID NO:499
Vinexin (SCAM-1)	NM_005775	SEQ ID NO:519	NP_005766	SEQ ID NO:520
Δ -6 fatty acid desaturase	NM_004265 AF134404	SEQ ID NO 521 SEQ ID NO:523	NP_004256 AAD31282	SEQ ID NO:522 SEQ ID NO:524

The terms "angiogenesis regulatory protein or nucleic acid" or a nucleic acid encoding "angiogenesis regulatory protein" refer to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino acid sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acids, to a polypeptide encoded by a nucleic acid of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, and SEQ ID NO:519, or an amino acid sequence of SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:520, SEQ ID NO:522, and SEQ ID NO:524 or other sequences listed herein; (2) specifically bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence of SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ

ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:520, SEQ ID NO:522, and SEQ ID NO:524, immunogenic fragments thereof, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid encoding

10 SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, 15 SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:520, SEQ ID NO:522, and SEQ ID NO:524, e.g., a nucleic 20 acid sequence of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID 25 NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID 30 NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523; or its complement, and conservatively modified variants thereof; (4) have a nucleic acid sequence that has greater than about 95%, preferably greater than about 96%, 97%, 98%, 99%, or higher nucleotide

sequence identity, preferably over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523, or the complement of any of those sequences. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or any mammal. The nucleic acids and proteins of the invention include both naturally occurring or recombinant molecules. When available, accession numbers for the human angiogenesis regulatory proteins and genes are provided.

The phrase "functional effects" in the context of assays for testing compounds that modulate activity of an angiogenesis regulatory proteins includes the determination of a parameter that is indirectly or directly under the influence of an angiogenesis regulatory protein polypeptide, e.g., a chemical or phenotypic effect such as loss-of angiogenesis phenotype represented by a change in expression of a cell surface marker, such as $\alpha v \beta 3$ integrin, or changes in cellular proliferation, especially endothelial cell proliferation; or enzymatic activity, or, e.g., a physical effect such as ligand binding or inhibition of ligand binding. A functional effect therefore includes ligand binding activity, the ability of cells to proliferate, expression in cells undergoing angiogenesis, and other characteristics of angiogenic cells. "Functional effects" include *in vitro*, *in vivo*, and *ex vivo* activities. Angiogenesis assays are described, e.g., in *Angiogenesis Protocols* (Murray, ed., 2001).

By "determining the functional effect" is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of an angiogenesis regulatory proteins, e.g., measuring physical and chemical or phenotypic effects. Such functional effects can be measured by any means known to those skilled in the

art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index); hydrodynamic (e.g., shape), chromatographic; or solubility properties for the protein; measuring inducible markers or transcriptional activation of the angiogenesis protein; measuring changes in cell surface markers, e.g., $\alpha v \beta 3$ integrin; and measuring cellular proliferation, particularly endothelial cell proliferation. Functional effects can be measured using ligand binding assays, e.g., binding to antibodies or other molecules. Functional effects can be determined by measuring changes in enzymatic activity. Examples of enzymatic activities include, but are not limited to protein translation, electron transfer, fatty acid desaturase activity, protein disulfide isomerase activity, transcriptional activation activity, peroxidase activity, kinase activity, and proteolytic activity.

Determination of the functional effect of a compound on angiogenesis can also be performed using angiogenesis assays known to those of skill in the art such as $\alpha v \beta 3$ downregulation, VEGF-2 upregulation, endothelial cell tube formation assays; chemotaxis assays, haptotaxis assays; differentiation assays using matrigel or co-culture with smooth muscle cells, the chick CAM assay; the mouse corneal assay; and assays that assess vascularization of an implanted tumor. The functional effects can be evaluated by many means known to those skilled in the art, e.g., microscopy for quantitative or qualitative measures of alterations in morphological features, e.g., tube or blood vessel formation, measurement of changes in RNA or protein levels for angiogenesis-associated sequences, measurement of RNA stability, identification of downstream or reporter gene expression (CAT, luciferase, β -gal, GFP and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, etc.

“Ligand” refers to a molecule that is specifically bound by a protein. An antibody is one example of a ligand. Other examples of ligands include RNA, cytochrome C, O_2 , NADP or NADPH, heme, MHC molecules, an amino acid sequence comprising a nuclear localization signal, tubulin or other cytoskeletal proteins, the thrombin receptor, transcription factor 7, cadherins, and ssDNA.

“Substrate” refers to a molecule that binds to an enzyme and is part of a specific chemical reaction catalyzed by the enzyme. Substrates of angiogenesis regulatory proteins include ATP, fatty acids, GTP, O_2 , arachadonic acids or its precursors, collagen, and protein molecules or peptides.

“Inhibitors,” “activators,” and “modulators” of angiogenesis regulatory protein polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules identified using *in vitro* and *in vivo* assays of angiogenesis regulatory

protein polynucleotide and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of angiogenesis regulatory proteins, e.g., antagonists. "Activators" are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate angiogenesis regulatory proteins activity, e.g., agonists. Inhibitors, activators, or modulators also include genetically modified versions of angiogenesis regulatory proteins, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, peptides, cyclic peptides, nucleic acids, antisense molecules, ribozymes, RNAi, small organic molecules and the like. Such assays for inhibitors and activators include, e.g., expressing angiogenesis regulatory proteins *in vitro*, in cells, or cell extracts, applying putative modulator compounds, and then determining the functional effects on activity, as described above.

Samples or assays comprising angiogenesis regulatory proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of angiogenesis regulatory protein is achieved when the activity value relative to the control is about 80%, preferably 50%, more preferably 25-0%. Activation of angiogenesis regulatory protein is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes any molecule, either naturally occurring or synthetic, e.g., protein, oligopeptide (e.g., from about 5 to about 25 amino acids in length, preferably from about 10 to 20 or 12 to 18 amino acids in length, preferably 12, 15, or 18 amino acids in length), small organic molecule, polysaccharide, lipid, fatty acid, polynucleotide, oligonucleotide, etc., to be tested for the capacity to directly or indirectly modulate angiogenesis. The test compound can be in the form of a library of test compounds, such as a combinatorial or randomized library that provides a sufficient range of diversity. Test compounds are optionally linked to a fusion partner, e.g., targeting compounds, rescue compounds, dimerization compounds, stabilizing compounds, addressable compounds, and other functional moieties. Conventionally, new chemical entities with useful properties are generated by identifying a test compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and

evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

A "small organic molecule" refers to an organic molecule, either naturally occurring or synthetic, that has a molecular weight of more than about 50 daltons and less than about 2500 daltons, preferably less than about 2000 daltons, preferably between about 100 to about 1000 daltons, more preferably between about 200 to about 500 daltons.

"RNAi molecule" or an "siRNA" refers to a nucleic acid that forms a double stranded RNA, which double stranded RNA has the ability to reduce or inhibit expression of a gene or target gene when the siRNA expressed in the same cell as the gene or target gene.

"siRNA" thus refers to the double stranded RNA formed by the complementary strands. The complementary portions of the siRNA that hybridize to form the double stranded molecule typically have substantial or complete identity. In one embodiment, an siRNA refers to a nucleic acid that has substantial or complete identity to a target gene and forms a double stranded siRNA. The sequence of the siRNA can correspond to the full length target gene, or a subsequence thereof. Typically, the siRNA is at least about 15-50 nucleotides in length (e.g., each complementary sequence of the double stranded siRNA is 15-50 nucleotides in length, and the double stranded siRNA is about 15-50 base pairs in length, preferable about 20-30 base nucleotides, preferably about 20-25 nucleotides in length, e.g., 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length.

"Biological sample" include sections of tissues such as biopsy and autopsy samples, and frozen sections taken for histologic purposes. Such samples include blood, sputum, tissue, cultured cells, e.g., primary cultures, explants, and transformed cells, stool, urine, etc. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate, e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or a bird; reptile; or fish.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 70% identity, preferably 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region (e.g., SEQ ID NO:1 or 2), when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (see, e.g., NCBI web site or the like). Such sequences are then said to be

“substantially identical.” This definition also refers to, or may be applied to, the compliment of a test sequence. The definition also includes sequences that have deletions and/or additions, as well as those that have substitutions. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A “comparison window”, as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

A preferred example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*, *Nuc. Acids Res.* 25:3389-3402 (1977) and Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990), respectively. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>).

This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*).

5 These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino
10 acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W , T , and X determine the
15 sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, $M=5$, $N=-4$ and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of
20 50, expectation (E) of 10, $M=5$, $N=-4$, and a comparison of both strands.

"Nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in single- or double-stranded form, or complements thereof. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have
25 similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs). Nucleic acids also include complementary nucleic acids.

30 Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or

deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* 260:2605-2608 (1985); Rossolini *et al.*, *Mol. Cell. Probes* 8:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide, and polynucleotide.

5 A particular nucleic acid sequence also implicitly encompasses "splice variants." Similarly, a particular protein encoded by a nucleic acid implicitly encompasses any protein encoded by a splice variant of that nucleic acid. "Splice variants," as the name suggests, are products of alternative splicing of a gene. After transcription, an initial nucleic acid transcript may be spliced such that different (alternate) nucleic acid splice products
10 encode different polypeptides. Mechanisms for the production of splice variants vary, but include alternate splicing of exons. Alternate polypeptides derived from the same nucleic acid by read-through transcription are also encompassed by this definition. Any products of a splicing reaction, including recombinant forms of the splice products, are included in this definition.

15 The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

20 The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have
25 the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical
30 compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

 Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical

Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

“Conservatively modified variants” applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are “silent variations,” which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence with respect to the expression product, but not with respect to actual probe sequences.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a “conservatively modified variant” where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

The following eight groups each contain amino acids that are conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (*see, e.g., Creighton, Proteins* (1984)).

Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, *see, e.g., Alberts et al., Molecular Biology of the Cell* (3rd ed., 1994) and Cantor and Schimmel, *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules* (1980).

5 “Primary structure” refers to the amino acid sequence of a particular peptide. “Secondary structure” refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains, e.g., transmembrane domains, pore domains, and cytoplasmic tail domains. Domains are portions of a polypeptide that form a compact unit of the polypeptide and are typically 15 to 350 amino acids long. Exemplary domains
10 include domains with enzymatic activity, e.g., phosphatase domains, ligand binding domains, etc. Typical domains are made up of sections of lesser organization such as stretches of β -sheet and α -helices. “Tertiary structure” refers to the complete three dimensional structure of a polypeptide monomer. “Quaternary structure” refers to the three dimensional structure formed by the noncovalent association of independent tertiary units. Anisotropic terms are
15 also known as energy terms.

A “label” or a “detectable moiety” is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, chemical, or other physical means. For example, useful labels include ³²P, fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins
20 which can be made detectable, e.g., by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide.

The term “recombinant” when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a
25 native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

The term “heterologous” when used with reference to portions of a nucleic
30 acid indicates that the nucleic acid comprises two or more subsequences that are not found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein indicates that the protein comprises

two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42°C, or, 5x SSC, 1% SDS, incubating at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., and *Current Protocols in Molecular Biology*, ed. Ausubel, *et al.*

For PCR, a temperature of about 36°C is typical for low stringency amplification, although annealing temperatures may vary between about 32°C and 48°C

depending on primer length. For high stringency PCR amplification, a temperature of about 62°C is typical, although high stringency annealing temperatures can range from about 50°C to about 65°C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90°C - 95°C for 30 sec - 2 min., an annealing phase lasting 30 sec. - 2 min., and an extension phase of about 72°C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.).

“Antibody” refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. Typically, the antigen-binding region of an antibody will be most critical in specificity and affinity of binding.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one “light” (about 25 kD) and one “heavy” chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce $F(ab)'_2$, a dimer of Fab which itself is a light chain joined to V_H-C_H1 by a disulfide bond. The $F(ab)'_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the $F(ab)'_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (*see Fundamental Immunology* (Paul ed., 3d ed. 1993)). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (*see, e.g., McCafferty et al., Nature* 348:552-554 (1990))

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many techniques known in the art can be used (*see, e.g., Kohler & Milstein, Nature* 256:495-497 (1975); Kozbor *et al., Immunology Today* 4: 72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985); Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *Antibodies, A Laboratory Manual* (1988); and Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986)). The genes encoding the heavy and light chains of an antibody of interest can be cloned from a cell, e.g., the genes encoding a monoclonal antibody can be cloned from a hybridoma and used to produce a recombinant monoclonal antibody. Gene libraries encoding heavy and light chains of monoclonal antibodies can also be made from hybridoma or plasma cells. Random combinations of the heavy and light chain gene products generate a large pool of antibodies with different antigenic specificity (*see, e.g., Kuby, Immunology* (3rd ed. 1997)). Techniques for the production of single chain antibodies or recombinant antibodies (U.S. Patent 4,946,778, U.S. Patent No. 4,816,567) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized or human antibodies (*see, e.g., U.S. Patent* Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, Marks *et al., Bio/Technology* 10:779-783 (1992); Lonberg *et al., Nature* 368:856-859 (1994); Morrison, *Nature* 368:812-13 (1994); Fishwild *et al., Nature Biotechnology* 14:845-51 (1996); Neuberger, *Nature Biotechnology* 14:826 (1996); and Lonberg & Huszar, *Intern. Rev. Immunol.* 13:65-93 (1995)). Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (*see, e.g., McCafferty et al., Nature* 348:552-554 (1990); Marks *et al., Biotechnology* 10:779-783 (1992)). Antibodies can also be made bispecific, i.e., able to recognize two different antigens (*see, e.g., WO 93/08829, Traunecker et al., EMBO J.* 10:3655-3659 (1991); and Suresh *et al., Methods in Enzymology* 121:210 (1986)). Antibodies can also be heteroconjugates, e.g., two covalently joined antibodies, or immunotoxins (*see, e.g., U.S. Patent No. 4,676,980, WO* 91/00360; WO 92/200373; and EP 03089).

Methods for humanizing or primatizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues

are often referred to as import residues, which are typically taken from an import variable domain. Humanization can be essentially performed following the method of Winter and co-workers (*see, e.g., Jones et al., Nature* 321:522-525 (1986); Riechmann *et al., Nature* 332:323-327 (1988); Verhoeyen *et al., Science* 239:1534-1536 (1988) and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, *e.g., an enzyme, toxin, hormone, growth factor, drug, etc.*; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

In one embodiment, the antibody is conjugated to an "effector" moiety. The effector moiety can be any number of molecules, including labeling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the antibody modulates the activity of the protein.

The phrase "specifically (or selectively) binds" to an antibody or "specifically (or selectively) immunoreactive with," when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein, often in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein at least two times the background and more typically more than 10 to 100 times background. Specific binding to an antibody under such conditions requires an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to angiogenesis regulatory proteins, polymorphic variants, alleles, orthologs, and conservatively modified variants, or splice variants, or portions thereof, can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with angiogenesis regulatory proteins and not with other proteins. This selection may be achieved by subtracting out antibodies that cross-react with

other molecules. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (*see, e.g., Harlow & Lane, Antibodies, A Laboratory Manual* (1988) for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity).

Assays for proteins that modulate angiogenesis

High throughput functional genomics assays can be used to identify modulators of angiogenesis. Such assays can monitor changes in cell surface marker expression, $\alpha v \beta 3$ integrin production, proliferation, and differentiation using either cell lines or primary cells, e.g., endothelial cells such as HUVEC cells. Typically, early passage or primary endothelial cells are contacted with a cDNA or a random peptide library (encoded by nucleic acids). Often the libraries are retroviral libraries. The cDNA library can comprise sense, antisense, full length, and truncated cDNAs. The cDNAs are often from endothelial cells, e.g., HUVEC cells, either differentiated on matrigel or 3-d collagen, or undifferentiated. The peptide library is encoded by nucleic acids. The cDNAs are optionally fused to markers such as GFP. The effect of the cDNA or peptide library on the endothelial cells is then monitored, using an assay such as cell surface marker expression (e.g., $\alpha v \beta 3$ integrin or VEGF2-R2) or a phenotypic assay for angiogenesis such as migration towards an ECM (extracellular matrix) component (*see, e.g., Klenke et al., J. Cell Biol.* 4:961-972 (1998)) or endothelial cell tube formation assays, as well as other bioassays such as the chick CAM assay, the mouse corneal assay, haptotaxis assays, and assays measuring the effect of administering potential modulators on implanted tumors. The effect of the cDNA or peptide can be validated and distinguished from somatic mutations, using, e.g., regulatable expression of the nucleic acid such as expression from a tetracycline promoter. cDNAs and nucleic acids encoding peptides can be rescued using techniques known to those of skill in the art, e.g., using a sequence tags.

Proteins interacting with the peptide or with the protein encoded by the cDNA (e.g., angiogenesis regulatory protein) can be isolated using a yeast two-hybrid system, mammalian two hybrid system, immunoprecipitation and proteomics, phage display screen, etc. Targets so identified can be further used as bait in these assays to identify additional members of the angiogenesis pathway, which members are also targets for drug development (*see, e.g., Fields et al., Nature* 340:245 (1989); Vasavada et al., *Proc. Nat'l Acad. Sci. USA*

88:10686 (1991); Fearon *et al.*, *Proc. Nat'l Acad. Sci. USA* 89:7958 (1992); Dang *et al.*, *Mol. Cell. Biol.* 11:954 (1991); Chien *et al.*, *Proc. Nat'l Acad. Sci. USA* 9578 (1991); and U.S. Patent Nos. 5,283,173, 5,667,973, 5,468,614, 5,525,490, and 5,637,463).

Suitable endothelial cell lines include human umbilical vein cells (*see, e.g.*,
5 Jaffe *et al.*, *J. Clin. Invest.* 52:2745-2754 (1973)); human adult dermal capillary-derived cells (*see, e.g.*, Davison *et al.*, *In Vitro* 19:937-945 (1983)); human adipose capillary derived cells (*see, e.g.*, Kern *et al.*, *J. Clin. Invest.* 71:1822-1829 (1983); bovine aorta (*see, e.g.*, Booyse *et al.*, *Thromb. Diathes. Aemorrh.* 34:825-839 (1975); and rat brain capillary derived cells (*see, e.g.*, Bowman *et al.*, *In Vitro* 17:353-362 (1981)). For culture of endothelial cell lines,
10 explants, and primary cells, *see* Freshney *et al.*, *Culture of Animal Cells* (3rd ed. 1994). Suitable angiogenesis cell surface markers include $\alpha v \beta 3$ integrin (*see, e.g.*, Elicerir & Cheresh, *Cancer J. Sci. Am.* 6 Supp. 3:S245-249 (2000), Maeshima *et al.*, *J. Biol. Chem.* (June 8, 2001)).

Cell surface markers such as $\alpha v \beta 3$ can be assayed using fluorescently labeled
15 antibodies and FACS. Cell proliferation can be measured using ³H-thymidine or dye inclusion. Angiogenesis phenotype is measured by loss of phenotype observation.

cDNA libraries are made from any suitable source, preferably from endothelial cells. Libraries encoding random peptides are made according to techniques well known to those of skill in the art (*see, e.g.*, U.S. Patent No. 6,153,380, 6,114,111, and
20 6,180,343). Any suitable vector can be used for the cDNA and peptide libraries, including, *e.g.*, retroviral vectors.

In a preferred embodiment, target proteins that modulate angiogenesis or nucleic acids that encode the target proteins are identified using a high throughput cell based assay (using a microtiter plate format) and FACS screening for $\alpha v \beta 3$ cell surface expression.
25 cDNA libraries are made which include, *e.g.*, sense, antisense, full length, and truncated cDNAs. The cDNAs are cloned into a retroviral vector. Endothelial cells are infected with the library, cultured for a short effector phase and then the cells with reduced $\alpha v \beta 3$ surface levels are enriched by antibody staining and magnetic cell sorting. The enriched cell population is then sorted into microtiter plates using fluorescent antibodies and FACS.
30 Resultant cell colonies are analyzed by immunofluorescence for reduced $\alpha v \beta 3$ surface levels. Selected colonies are infected with wild type MMLV virus to rescue the proviral vector. The infectious supernatant is used to infect endothelial cells, which are subsequently analyzed for $\alpha v \beta 3$ levels by FACS. The cDNA is isolated and sequenced to determine if it represents a wild type or mutated cDNA, *e.g.*, whether the cDNA represents a negative transdominant

mutant. Optionally, a marker such as GFP can be used to select for retrovirally infected cells. Using this system, cDNA molecules encoding angiogenesis regulatory protein were identified as a targets for anti-angiogenic drug therapy. Some of the cDNAs were fused in frame to a GFP marker. Some cDNA molecules were expressed as inhibitory antisense molecules.

5 Finally, some cDNAs were expressed in the antisense orientation and expression of the translated protein or peptide resulted in an angiogenesis phenotype.

Isolation of nucleic acids encoding angiogenesis regulatory proteins

This invention relies on routine techniques in the field of recombinant
10 genetics. Basic texts disclosing the general methods of use in this invention include Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2nd ed. 1989); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); and *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., 1994)).

Nucleic acids encoding angiogenesis regulatory proteins and their
15 polymorphic variants, orthologs, and alleles that are substantially identical to the amino acid sequence of SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID
20 NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID
25 NO:497, SEQ ID NO:499, and SEQ ID NO:520 can be isolated using nucleic acid probes and oligonucleotides under stringent hybridization conditions, by screening libraries.

Alternatively, expression libraries can be used to clone angiogenesis regulatory proteins, polymorphic variants, orthologs, and alleles by detecting expressed homologs immunologically with antisera or purified antibodies made against a human angiogenesis
30 regulatory protein or portions thereof.

To make a cDNA library, one should choose a source that is rich in angiogenesis regulatory protein RNA, e.g., endothelial cells. The mRNA is then made into cDNA using reverse transcriptase, ligated into a recombinant vector, and transfected into a recombinant host for propagation, screening and cloning. Methods for making and screening

cDNA libraries are well known (*see, e.g.,* Gubler & Hoffman, *Gene* 25:263-269 (1983); Sambrook *et al., supra*; Ausubel *et al., supra*).

For a genomic library, the DNA is extracted from the tissue and either mechanically sheared or enzymatically digested to yield fragments of about 12-20 kb. The fragments are then separated by gradient centrifugation from undesired sizes and are constructed in bacteriophage lambda vectors. These vectors and phage are packaged *in vitro*. Recombinant phage are analyzed by plaque hybridization as described in Benton & Davis, *Science* 196:180-182 (1977). Colony hybridization is carried out as generally described in Grunstein *et al., Proc. Natl. Acad. Sci. USA.*, 72:3961-3965 (1975).

An alternative method of isolating angiogenesis regulatory protein nucleic acid and its orthologs, alleles, mutants, polymorphic variants, and conservatively modified variants combines the use of synthetic oligonucleotide primers and amplification of an RNA or DNA template (*see* U.S. Patents 4,683,195 and 4,683,202; *PCR Protocols: A Guide to Methods and Applications* (Innis *et al.*, eds, 1990)). Methods such as polymerase chain reaction (PCR) and ligase chain reaction (LCR) can be used to amplify nucleic acid sequences encoding human angiogenesis regulatory proteins directly from mRNA, from cDNA, from genomic libraries or cDNA libraries. Degenerate oligonucleotides can be designed to amplify angiogenesis regulatory protein homologs using the sequences provided herein. Restriction endonuclease sites can be incorporated into the primers. Polymerase chain reaction or other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of angiogenesis regulatory protein encoding mRNA in physiological samples, for nucleic acid sequencing, or for other purposes. Genes amplified by the PCR reaction can be purified from agarose gels and cloned into an appropriate vector.

Gene expression of angiogenesis regulatory proteins can also be analyzed by techniques known in the art, e.g., reverse transcription and amplification of mRNA, isolation of total RNA or poly A⁺ RNA, northern blotting, dot blotting, *in situ* hybridization, RNase protection, high density polynucleotide array technology, e.g., and the like.

Nucleic acids encoding angiogenesis regulatory proteins can be used with high density oligonucleotide array technology (e.g., GeneChipTM) to identify angiogenesis regulatory proteins, orthologs, alleles, conservatively modified variants, and polymorphic variants in this invention. In the case where the homologs being identified are linked to a condition such as neovascularization or angiogenic insufficiency, they can be used with GeneChipTM as a diagnostic tool in detecting the disease in a biological sample, *see, e.g.,*

Gunthand *et al.*, *AIDS Res. Hum. Retroviruses* 14: 869-876 (1998); Kozal *et al.*, *Nat. Med.* 2:753-759 (1996); Matson *et al.*, *Anal. Biochem.* 224:110-106 (1995); Lockhart *et al.*, *Nat. Biotechnol.* 14:1675-1680 (1996); Gingeras *et al.*, *Genome Res.* 8:435-448 (1998); Hacia *et al.*, *Nucleic Acids Res.* 26:3865-3866 (1998).

5 The gene for an angiogenesis regulatory protein is typically cloned into intermediate vectors before transformation into prokaryotic or eukaryotic cells for replication and/or expression. These intermediate vectors are typically prokaryote vectors, e.g., plasmids, or shuttle vectors.

10 **Expression in prokaryotes and eukaryotes**

To obtain high level expression of a cloned gene, such as those cDNAs encoding an angiogenesis regulatory protein, one typically subclones a gene encoding an angiogenesis regulatory protein into an expression vector that contains a strong promoter to direct transcription, a transcription/translation terminator, and if for a nucleic acid encoding a protein, a ribosome binding site for translational initiation. Suitable bacterial promoters are well known in the art and described, e.g., in Sambrook *et al.*, and Ausubel *et al.*, *supra*. Bacterial expression systems for expressing angiogenesis regulatory proteins are available in, e.g., *E. coli*, *Bacillus sp.*, and *Salmonella* (Palva *et al.*, *Gene* 22:229-235 (1983); Mosbach *et al.*, *Nature* 302:543-545 (1983). Kits for such expression systems are commercially available. Eukaryotic expression systems for mammalian cells, yeast, and insect cells are well known in the art and are also commercially available.

Selection of the promoter used to direct expression of a heterologous nucleic acid depends on the particular application. The promoter is preferably positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

In addition to the promoter, the expression vector typically contains a transcription unit or expression cassette that contains all the additional elements required for the expression of the angiogenesis regulatory protein encoding nucleic acid in host cells. A typical expression cassette thus contains a promoter operably linked to the nucleic acid sequence encoding an angiogenesis regulatory protein and signals required for efficient polyadenylation of the transcript, ribosome binding sites, and translation termination. Additional elements of the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns with functional splice donor and acceptor sites.

In addition to a promoter sequence, the expression cassette should also contain a transcription termination region downstream of the structural gene to provide for efficient termination. The termination region may be obtained from the same gene as the promoter sequence or may be obtained from different genes.

5 The particular expression vector used to transport the genetic information into the cell is not particularly critical. Any of the conventional vectors used for expression in eukaryotic or prokaryotic cells may be used. Standard bacterial expression vectors include plasmids such as pBR322 based plasmids, pSKF, pET23D, and fusion expression systems such as MBP, GST, and LacZ. Epitope tags can also be added to recombinant proteins to
10 provide convenient methods of isolation, e.g., c-myc.

Expression vectors containing regulatory elements from eukaryotic viruses are typically used in eukaryotic expression vectors, e.g., SV40 vectors, papilloma virus vectors, retroviral vectors, and vectors derived from Epstein-Barr virus. Other exemplary eukaryotic vectors include pMSG, pAV009/A⁺, pMTO10/A⁺, pMAMneo-5, baculovirus pDSVE, and
15 any other vector allowing expression of proteins under the direction of the CMV promoter, SV40 early promoter, SV40 later promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, polyhedrin promoter, or other promoters shown effective for expression in eukaryotic cells.

Expression of proteins from eukaryotic vectors can be also be regulated using
20 inducible promoters. With inducible promoters, expression levels are tied to the concentration of inducing agents, such as tetracycline or ecdysone, by the incorporation of response elements for these agents into the promoter. Generally, high level expression is obtained from inducible promoters only in the presence of the inducing agent; basal expression levels are minimal. Inducible expression vectors are often chosen if expression of
25 the protein of interest is detrimental to eukaryotic cells.

In one embodiment, the vectors of the invention have a regulatable promoter, e.g., tet-regulated systems and the RU-486 system (*see, e.g., Gossen & Bujard, Proc. Nat'l Acad. Sci. USA* 89:5547 (1992); Oligino *et al., Gene Ther.* 5:491-496 (1998); Wang *et al., Gene Ther.* 4:432-441 (1997); Neering *et al., Blood* 88:1147-1155 (1996); and Rendahl *et al.,*
30 *Nat. Biotechnol.* 16:757-761 (1998)). These impart small molecule control on the expression of the candidate target nucleic acids. This beneficial feature can be used to determine that a desired phenotype is caused by a transfected cDNA rather than a somatic mutation.

Some expression systems have markers that provide gene amplification such as thymidine kinase and dihydrofolate reductase. Alternatively, high yield expression

systems not involving gene amplification are also suitable, such as using a baculovirus vector in insect cells, with an angiogenesis regulatory protein encoding sequence under the direction of the polyhedrin promoter or other strong baculovirus promoters.

The elements that are typically included in expression vectors also include a replicon that functions in *E. coli*, a gene encoding antibiotic resistance to permit selection of bacteria that harbor recombinant plasmids, and unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. The particular antibiotic resistance gene chosen is not critical, any of the many resistance genes known in the art are suitable. The prokaryotic sequences are preferably chosen such that they do not interfere with the replication of the DNA in eukaryotic cells, if necessary.

Standard transfection methods are used to produce bacterial, mammalian, yeast or insect cell lines that express large quantities of angiogenesis regulatory proteins, which are then purified using standard techniques (*see, e.g., Colley et al., J. Biol. Chem.* 264:17619-17622 (1989); *Guide to Protein Purification*, in *Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). Transformation of eukaryotic and prokaryotic cells are performed according to standard techniques (*see, e.g., Morrison, J. Bact.* 132:349-351 (1977); Clark-Curtiss & Curtiss, *Methods in Enzymology* 101:347-362 (Wu *et al.*, eds, 1983).

Any of the well-known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, polybrene, protoplast fusion, electroporation, biolistics, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g., Sambrook et al., supra*). It is only necessary that the particular genetic engineering procedure used be capable of successfully introducing at least one gene into the host cell capable of expressing angiogenesis regulatory protein.

After the expression vector is introduced into the cells, the transfected cells are cultured under conditions favoring expression of angiogenesis regulatory protein, which is recovered from the culture using standard techniques identified below.

Purification of angiogenesis regulatory proteins

Either naturally occurring or recombinant angiogenesis regulatory proteins can be purified for use in functional assays. Naturally occurring angiogenesis regulatory protein can be purified, *e.g.,* from human tissue. Recombinant angiogenesis regulatory protein can be purified from any suitable expression system.

The angiogenesis regulatory proteins may be purified to substantial purity by standard techniques, including selective precipitation with such substances as ammonium sulfate; column chromatography, immunopurification methods, and others (*see, e.g., Scopes, Protein Purification: Principles and Practice* (1982); U.S. Patent No. 4,673,641; Ausubel *et al., supra*; and Sambrook *et al., supra*).

A number of procedures can be employed when recombinant angiogenesis regulatory proteins are being purified. For example, proteins having established molecular adhesion properties can be reversibly fused to the angiogenesis regulatory proteins. With the appropriate ligand, the angiogenesis regulatory proteins can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form. The fused protein is then removed by enzymatic activity. Finally, the angiogenesis regulatory proteins could be purified using immunoaffinity columns.

A. Purification of angiogenesis regulatory proteins from recombinant bacteria

Recombinant proteins are expressed by transformed bacteria in large amounts, typically after promoter induction; but expression can be constitutive. Promoter induction with IPTG is one example of an inducible promoter system. Bacteria are grown according to standard procedures in the art. Fresh or frozen bacteria cells are used for isolation of protein.

Proteins expressed in bacteria may form insoluble aggregates ("inclusion bodies"). Several protocols are suitable for purification of inclusion bodies containing angiogenesis regulatory proteins. For example, purification of inclusion bodies typically involves the extraction, separation and/or purification of inclusion bodies by disruption of bacterial cells, *e.g.,* by incubation in a buffer of 50 mM TRIS/HCL pH 7.5, 50 mM NaCl, 5 mM MgCl₂, 1 mM DTT, 0.1 mM ATP, and 1 mM PMSF. The cell suspension can be lysed using 2-3 passages through a French Press, homogenized using a Polytron (Brinkman Instruments) or sonicated on ice. Alternate methods of lysing bacteria are apparent to those of skill in the art (*see, e.g., Sambrook et al., supra; Ausubel et al., supra*).

If necessary, the inclusion bodies are solubilized, and the lysed cell suspension is typically centrifuged to remove unwanted insoluble matter. Proteins that formed the inclusion bodies may be renatured by dilution or dialysis with a compatible buffer. Suitable solvents include, but are not limited to urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to about 8 M). Some solvents which are capable of solubilizing aggregate-forming proteins, for example SDS (sodium dodecyl sulfate), 70% formic acid, are inappropriate for use in this

procedure due to the possibility of irreversible denaturation of the proteins, accompanied by a lack of immunogenicity and/or activity. Although guanidine hydrochloride and similar agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of immunologically and/or biologically active protein. Other suitable buffers are known to those skilled in the art. Human angiogenesis regulatory proteins are separated from other bacterial proteins by standard separation techniques, e.g., with Ni-NTA agarose resin.

Alternatively, it is possible to purify angiogenesis regulatory proteins from bacteria periplasm. After lysis of the bacteria, when the angiogenesis regulatory proteins are exported into the periplasm of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to skill in the art. To isolate recombinant proteins from the periplasm, the bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a buffer containing 20% sucrose. To lyse the cells, the bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM MgSO_4 and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of skill in the art.

B. Standard protein separation techniques for purifying angiogenesis regulatory proteins

Solubility fractionation

Often as an initial step, particularly if the protein mixture is complex, an initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest. The preferred salt is ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol includes adding saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This concentration will precipitate the most hydrophobic of proteins. The precipitate is then discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if

necessary, either through dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

Size differential filtration

The molecular weight of the angiogenesis regulatory proteins can be used to isolate it from proteins of greater and lesser size using ultrafiltration through membranes of different pore size (for example, Amicon or Millipore membranes). As a first step, the protein mixture is ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of the protein of interest. The retentate of the ultrafiltration is then ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

Column chromatography

The angiogenesis regulatory proteins can also be separated from other proteins on the basis of its size, net surface charge, hydrophobicity, and affinity for ligands. In addition, antibodies raised against proteins can be conjugated to column matrices and the proteins immunopurified. All of these methods are well known in the art. It will be apparent to one of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (e.g., Pharmacia Biotech).

Assays for modulators of angiogenesis regulatory proteins and angiogenesis

A. Assays

Modulation of a angiogenesis regulatory proteins, and corresponding modulation of angiogenesis, can be assessed using a variety of *in vitro* and *in vivo* assays, including high throughput ligand binding and cell based assays, as described herein. Such assays can be used to test for inhibitors and activators of angiogenesis regulatory proteins, and, consequently, inhibitors and activators of angiogenesis. Such modulators of angiogenesis regulatory proteins are useful for treating angiogenesis disorders. Modulators of angiogenesis regulatory proteins are tested using either recombinant or naturally occurring angiogenesis regulatory protein, preferably human angiogenesis regulatory protein.

Preferably, the angiogenesis regulatory proteins will have the sequence displayed in SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:520, SEQ ID NO:522, and SEQ ID NO:524 or a conservatively modified variant thereof. Alternatively, the angiogenesis regulatory proteins of the assay will be derived from a eukaryote and include an amino acid subsequence having substantial amino acid sequence identity to SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:520, SEQ ID NO:522, and SEQ ID NO:524. Generally, the amino acid sequence identity will be at least 60%, preferably at least 65%, 70%, 75%, 80%, 85%, or 90%, most preferably at least 95%.

Measurement of an angiogenic or loss-of-angiogenesis phenotype in angiogenesis regulatory proteins or cells expressing angiogenesis regulatory proteins, either recombinant or naturally occurring, can be performed using a variety of assays, *in vitro*, *in vivo*, and *ex vivo*. For example, recombinant or naturally occurring angiogenesis regulatory protein can be used *in vitro*, in a ligand binding or enzymatic function assay. Angiogenesis regulatory proteins present in a cellular extract can also be used in *in vitro* assays. Cell- and animal-based *in vivo* assays can also be used to assay for angiogenesis regulatory protein modulators. Any suitable physical, chemical, or phenotypic change that affects activity or binding can be used to assess the influence of a test compound on the polypeptide of this

invention. When the functional effects are determined using intact cells or animals, one can also measure a variety of effects such as, in the case of angiogenesis associated with tumors, tumor growth, neovascularization, cell surface markers such as $\alpha v \beta 3$, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as cGMP. In one embodiment, measurement of $\alpha v \beta 3$ integrin cell surface expression and FACS sorting is used to identify modulators of angiogenesis.

In vitro assays

Assays to identify compounds with angiogenesis regulatory protein modulating activity, e.g., anti-angiogenic activity, can be performed *in vitro*, e.g., binding assays. Such assays can use full length angiogenesis regulatory proteins or a variant thereof (see, e.g., SEQ ID NO:2), or a fragment of an angiogenesis regulatory proteins having a desired activity. Purified recombinant or naturally occurring angiogenesis regulatory proteins can be used in the *in vitro* methods of the invention. In addition to purified angiogenesis regulatory proteins, the recombinant or naturally occurring angiogenesis regulatory proteins can be part of a cellular lysate. As described below, the assay can be either solid state or soluble. Preferably, the protein is bound to a solid support, either covalently or non-covalently. Often, the *in vitro* assays of the invention are ligand binding or ligand affinity assays, either non-competitive or competitive. Other *in vitro* assays include measuring changes in spectroscopic (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein.

In one embodiment, a high throughput binding assay is performed in which the angiogenesis regulatory proteins or chimera comprising a fragment thereof is contacted with a potential modulator and incubated for a suitable amount of time. In one embodiment, the potential modulator is bound to a solid support, and the angiogenesis regulatory proteins is added. In another embodiment, the angiogenesis regulatory proteins is bound to a solid support. A wide variety of modulators can be used, as described below, including small organic molecules, peptides, and antibodies. A wide variety of assays can be used to identify angiogenesis regulatory protein -modulator binding or enzymatic activity, including labeled protein-protein binding assays, electrophoretic mobility shifts, immunoassays, and the like. In some cases, the binding of the candidate modulator is determined through the use of competitive binding assays, where interference with binding of a known ligand is measured

in the presence of a potential modulator. Often, either the potential modulator or the known ligand is labeled. Binding of the candidate modulator can also be assessed by determining its effect on the enzymatic activity of an angiogenesis regulatory protein.

5 Cell-based *in vivo* assays

In another embodiment, an angiogenesis regulatory proteins is expressed in a cell, and functional, e.g., physical and chemical or phenotypic, changes are assayed to identify angiogenesis modulators, preferably anti-angiogenesis compounds. Cells expressing angiogenesis regulatory proteins can also be used in binding assays or enzymatic assays.

10 Any suitable functional effect can be measured, as described herein. For example, ligand binding, cell surface marker expression, cellular proliferation, cellular differentiation assays and cell migration assays are all suitable assays to identify potential modulators using a cell based system. Suitable cells for such cell based assays include both primary endothelial cells and cell lines, as described herein. The angiogenesis regulatory proteins can be naturally
15 occurring or recombinant. Also, as described above, a fragment of an angiogenesis regulatory protein with a desired activity can be used in cell based assays.

As described above, in one embodiment, loss-of angiogenesis phenotype is measured by contacting endothelial cells comprising an angiogenesis regulatory protein target with a potential modulator. Modulation of angiogenesis is identified by screening for cell
20 surface marker expression, e.g., $\alpha v\beta 3$ integrin expression levels, using fluorescent antibodies and FACS sorting.

In another embodiment, cellular angiogenesis regulatory protein levels are determined by measuring the level of protein or mRNA. The level of angiogenesis regulatory protein or proteins are measured using immunoassays such as western blotting, ELISA and
25 the like with an antibody that selectively binds to the angiogenesis regulatory protein polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNase protection, dot blotting, are preferred. The level of protein or mRNA is detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids,
30 radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, angiogenesis regulatory protein expression can be measured using a reporter gene system. Such a system can be devised using a angiogenesis regulatory proteins promoter operably linked to a reporter gene such as chloramphenicol

acetyltransferase, firefly luciferase, bacterial luciferase, β -galactosidase and alkaline phosphatase. Furthermore, the protein of interest can be used as an indirect reporter via attachment to a second reporter such as red or green fluorescent protein (*see, e.g.,* Mistili & Spector, *Nature Biotechnology* 15:961-964 (1997)). The reporter construct is typically
5 transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art.

In another embodiment, angiogenesis regulatory protein enzymatic activity can be measured, using, e.g., labeled substrate proteins, gel electrophoresis, and ELISA
10 assays, and other appropriate enzymatic activities.

A variety of phenotypic angiogenesis assays are known to those of skill in the art. Various models have been employed to evaluate angiogenesis (e.g., Croix *et al.*, *Science* 289:1197-1202 (2000) and Kahn *et al.*, *Amer. J. Pathol.* 156:1887-1900). Assessment of angiogenesis in the presence of a potential modulator of angiogenesis can be performed using
15 cell-culture-based angiogenesis assays, e.g., endothelial cell tube formation assays, cellular differentiation assays using matrigel matrix or by co-culture with smooth muscle cells, chemotaxis assays using VEGF or FGF, and haptotaxis assays, as well as other animal based bioassays such as the chick CAM assay, the mouse corneal assay, and assays measuring the effect of administering potential modulators on implanted tumors.

Animal models

A number of animal based assays for angiogenesis phenotypes are known to those of skill in the art and can be used to assay for modulators of angiogenesis. For example, the chick CAM assay is described by O'Reilly, *et al.* *Cell* 79: 315-328 (1994).
25 Briefly, 3 day old chicken embryos with intact yolks are separated from the egg and placed in a petri dish. After 3 days of incubation, a methylcellulose disc containing the protein to be tested is applied to the CAM of individual embryos. After about 48 hours of incubation, the embryos and CAMs are observed to determine whether endothelial growth has been inhibited.

The mouse corneal assay involves implanting a growth factor-containing pellet, along with another pellet containing the suspected endothelial growth inhibitor, in the cornea of a mouse and observing the pattern of capillaries that are elaborated in the cornea.
30

Angiogenesis can also be measured by determining the extent of neovascularization of a tumor. For example, carcinoma cells can be subcutaneously inoculated into athymic nude mice and tumor growth then monitored. Immunoassays using endothelial cell-specific antibodies are typically used to stain for vascularization of tumor and the number of vessels in the tumor.

As described above, animal models of angiogenesis find use in screening for modulators of angiogenesis. Similarly, transgenic animal technology including gene knockout technology, for example as a result of homologous recombination with an appropriate gene targeting vector, or gene overexpression, will result in the absence or increased expression of the angiogenesis regulatory proteins. The same technology can also be applied to make knock-out cells. When desired, tissue-specific expression or knockout of the gene encoding an angiogenesis regulatory protein may be necessary. Transgenic animals generated by such methods find use as animal models of angiogenesis and are additionally useful in screening for modulators of angiogenesis.

Knock-out cells and transgenic mice can be made by insertion of a marker gene or other heterologous gene into the endogenous gene site encoding an angiogenesis regulatory protein in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous angiogenesis regulatory protein with a mutated version of angiogenesis regulatory protein, or by mutating the endogenous angiogenesis regulatory protein, e.g., by exposure to carcinogens.

A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (*see, e.g., Capecchi et al., Science* 244:1288 (1989)). Chimeric targeted mice can be derived according to Hogan *et al., Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbor Laboratory (1988) and *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, ed., IRL Press, Washington, D.C., (1987).

B. Modulators

The compounds tested as modulators of angiogenesis regulatory proteins can be any small organic molecule, or a biological entity, such as a protein, e.g., an antibody or peptide, a sugar, a nucleic acid, e.g., an antisense oligonucleotide, RNAi, or a ribozyme, or a

lipid. Alternatively, modulators can be genetically altered versions of an angiogenesis regulatory proteins. Typically, test compounds will be small organic molecules, peptides, lipids, and lipid analogs.

Essentially any chemical compound can be used as a potential modulator or
5 ligand in the assays of the invention, although most often compounds can be dissolved in aqueous or organic (especially DMSO-based) solutions are used. The assays are designed to screen large chemical libraries by automating the assay steps and providing compounds from any convenient source to assays, which are typically run in parallel (e.g., in microtiter formats on microtiter plates in robotic assays). It will be appreciated that there are many suppliers of
10 chemical compounds, including Sigma (St. Louis, MO), Aldrich (St. Louis, MO), Sigma-Aldrich (St. Louis, MO), Fluka Chemika-Biochemica Analytika (Buchs Switzerland) and the like.

In one preferred embodiment, high throughput screening methods involve providing a combinatorial small organic molecule or peptide library containing a large
15 number of potential therapeutic compounds (potential modulator or ligand compounds). Such "combinatorial chemical libraries" or "ligand libraries" are then screened in one or more assays, as described herein, to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential
20 or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis, by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical
25 building blocks (amino acids) in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (*see, e.g.,* U.S. Patent 5,010,175, Furka, *Int. J. Pept. Prot. Res.* 37:487-493 (1991) and Houghton *et al., Nature* 354:84-88 (1991)). Other chemistries for generating chemical diversity libraries can also be used. Such chemistries include, but are not limited to: peptoids (e.g., PCT Publication No. WO 91/19735), encoded peptides (e.g., PCT Publication No. WO 93/20242), random bio-oligomers (e.g., PCT Publication No. WO 92/00091),

benzodiazepines (e.g., U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al.*, *Proc. Nat. Acad. Sci. USA* 90:6909-6913 (1993)), vinylogous polypeptides (Hagihara *et al.*, *J. Amer. Chem. Soc.* 114:6568 (1992)), nonpeptidal peptidomimetics with glucose scaffolding (Hirschmann *et al.*, *J. Amer. Chem. Soc.* 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen *et al.*, *J. Amer. Chem. Soc.* 116:2661 (1994)), oligocarbamates (Cho *et al.*, *Science* 261:1303 (1993)), and/or peptidyl phosphonates (Campbell *et al.*, *J. Org. Chem.* 59:658 (1994)), nucleic acid libraries (*see* Ausubel, Berger and Sambrook, all *supra*), peptide nucleic acid libraries (*see, e.g.*, U.S. Patent 5,539,083), antibody libraries (*see, e.g.*, Vaughn *et al.*, *Nature Biotechnology*, 14(3):309-314 (1996) and PCT/US96/10287), carbohydrate libraries (*see, e.g.*, Liang *et al.*, *Science*, 274:1520-1522 (1996) and U.S. Patent 5,593,853), small organic molecule libraries (*see, e.g.*, benzodiazepines, Baum C&EN, Jan 18, page 33 (1993); isoprenoids, U.S. Patent 5,569,588; thiazolidinones and metathiazanones, U.S. Patent 5,549,974; pyrrolidines, U.S. Patents 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent 5,506,337; benzodiazepines, 5,288,514, and the like).

Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.*, 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA). In addition, numerous combinatorial libraries are themselves commercially available (*see, e.g.*, ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, etc.).

C. Solid State and soluble high throughput assays

In one embodiment the invention provides soluble assays using angiogenesis regulatory proteins, or a cell or tissue expressing angiogenesis regulatory proteins, either naturally occurring or recombinant. In another embodiment, the invention provides solid phase based *in vitro* assays in a high throughput format, where the angiogenesis regulatory protein is attached to a solid phase substrate. Any one of the assays described herein can be adapted for high throughput screening, e.g., ligand binding, cellular proliferation, cell surface marker flux, e.g., $\alpha\beta 3$ integrin, enzymatic activity, etc.. In one preferred embodiment, the cell-based system using $\alpha\beta 3$ integrin modulation and FACS assays is used in a high throughput format for identifying modulators of angiogenesis regulatory proteins, and therefore modulators of T cell angiogenesis.

In the high throughput assays of the invention, either soluble or solid state, it is possible to screen up to several thousand different modulators or ligands in a single day. This methodology can be used for angiogenesis regulatory proteins *in vitro*, or for cell-based assays comprising an angiogenesis regulatory proteins. In particular, each well of a
5 microtiter plate can be used to run a separate assay against a selected potential modulator, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 100 (e.g., 96) modulators. If 1536 well plates are used, then a single plate can easily assay from about 100-
10 about 1500 different compounds. It is possible to assay many plates per day; assay screens for up to about 6,000, 20,000, 50,000, or more than 100,000 different compounds are possible using the integrated systems of the invention.

For a solid state reaction, the protein of interest or a fragment thereof, e.g., an extracellular domain, or a cell comprising the protein of interest or a fragment thereof as part of a fusion protein can be bound to the solid state component, directly or indirectly, via
15 covalent or non covalent linkage e.g., via a tag. The tag can be any of a variety of components. In general, a molecule which binds the tag (a tag binder) is fixed to a solid support, and the tagged molecule of interest is attached to the solid support by interaction of the tag and the tag binder.

A number of tags and tag binders can be used, based upon known molecular
20 interactions well described in the literature. For example, where a tag has a natural binder, for example, biotin, protein A, or protein G, it can be used in conjunction with appropriate tag binders (avidin, streptavidin, neutravidin, the Fc region of an immunoglobulin, etc.) Antibodies to molecules with natural binders such as biotin are also widely available and appropriate tag binders; *see*, SIGMA Immunochemicals 1998 catalogue SIGMA, St. Louis
25 MO).

Similarly, any haptenic or antigenic compound can be used in combination with an appropriate antibody to form a tag/tag binder pair. Thousands of specific antibodies are commercially available and many additional antibodies are described in the literature. For example, in one common configuration, the tag is a first antibody and the tag binder is a
30 second antibody which recognizes the first antibody. In addition to antibody-antigen interactions, receptor-ligand interactions are also appropriate as tag and tag-binder pairs. For example, agonists and antagonists of cell membrane receptors (e.g., cell receptor-ligand interactions such as transferrin, c-kit, viral receptor ligands, cytokine receptors, chemokine receptors, interleukin receptors, immunoglobulin receptors and antibodies, the cadherein

family, the integrin family, the selectin family, and the like; *see, e.g.*, Pigott & Power, *The Adhesion Molecule Facts Book I* (1993). Similarly, toxins and venoms, viral epitopes, hormones (e.g., opiates, steroids, etc.), intracellular receptors (e.g. which mediate the effects of various small ligands, including steroids, thyroid hormone, retinoids and vitamin D; peptides), drugs, lectins, sugars, nucleic acids (both linear and cyclic polymer configurations), oligosaccharides, proteins, phospholipids and antibodies can all interact with various cell receptors.

Synthetic polymers, such as polyurethanes, polyesters, polycarbonates, polyureas, polyamides, polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides, and polyacetates can also form an appropriate tag or tag binder. Many other tag/tag binder pairs are also useful in assay systems described herein, as would be apparent to one of skill upon review of this disclosure.

Common linkers such as peptides, polyethers, and the like can also serve as tags, and include polypeptide sequences, such as poly gly sequences of between about 5 and 200 amino acids. Such flexible linkers are known to persons of skill in the art. For example, poly(ethylene glycol) linkers are available from Shearwater Polymers, Inc. Huntsville, Alabama. These linkers optionally have amide linkages, sulfhydryl linkages, or heterofunctional linkages.

Tag binders are fixed to solid substrates using any of a variety of methods currently available. Solid substrates are commonly derivatized or functionalized by exposing all or a portion of the substrate to a chemical reagent which fixes a chemical group to the surface which is reactive with a portion of the tag binder. For example, groups which are suitable for attachment to a longer chain portion would include amines, hydroxyl, thiol, and carboxyl groups. Aminoalkylsilanes and hydroxyalkylsilanes can be used to functionalize a variety of surfaces, such as glass surfaces. The construction of such solid phase biopolymer arrays is well described in the literature. *See, e.g.*, Merrifield, *J. Am. Chem. Soc.* 85:2149-2154 (1963) (describing solid phase synthesis of, e.g., peptides); Geysen *et al.*, *J. Immun. Meth.* 102:259-274 (1987) (describing synthesis of solid phase components on pins); Frank & Doring, *Tetrahedron* 44:60316040 (1988) (describing synthesis of various peptide sequences on cellulose disks); Fodor *et al.*, *Science*, 251:767-777 (1991); Sheldon *et al.*, *Clinical Chemistry* 39(4):718-719 (1993); and Kozal *et al.*, *Nature Medicine* 2(7):753759 (1996) (all describing arrays of biopolymers fixed to solid substrates). Non-chemical approaches for fixing tag binders to substrates include other common methods, such as heat, cross-linking by UV radiation, and the like.

Antibodies to angiogenesis regulatory proteins

In addition to the detection of genes encoding angiogenesis regulatory proteins and gene expression using nucleic acid hybridization technology, one can also use

immunoassays to detect angiogenesis regulatory proteins of the invention. Such assays are useful for screening for modulators of angiogenesis regulatory proteins and angiogenesis, as well as for therapeutic and diagnostic applications. Immunoassays can be used qualitatively or quantitatively to analyze angiogenesis regulatory proteins. A general overview of the applicable technology can be found in Harlow & Lane, *Antibodies: A Laboratory Manual* (1988).

A. Production of antibodies

Methods of producing polyclonal and monoclonal antibodies that react specifically with the angiogenesis regulatory proteins are known to those of skill in the art (see, e.g., Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *supra*; Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986); and Kohler & Milstein, *Nature* 256:495-497 (1975). Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors, as well as preparation of polyclonal and monoclonal antibodies by immunizing rabbits or mice (see, e.g., Huse *et al.*, *Science* 246:1275-1281 (1989); Ward *et al.*, *Nature* 341:544-546 (1989)).

A number of immunogens comprising portions of angiogenesis regulatory proteins may be used to produce antibodies specifically reactive with angiogenesis regulatory proteins. For example, recombinant angiogenesis regulatory proteins or an antigenic fragment thereof, can be isolated as described herein. Recombinant protein can be expressed in eukaryotic or prokaryotic cells as described above, and purified as generally described above. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used as an immunogen. Naturally occurring protein may also be used either in pure or impure form. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the protein.

Methods of production of polyclonal antibodies are known to those of skill in the art. An inbred strain of mice (e.g., BALB/C mice) or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization

protocol. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the beta subunits. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (*see*, Harlow & Lane, *supra*).

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*see*, Kohler & Milstein, *Eur. J. Immunol.* 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse, *et al.*, *Science* 246:1275-1281 (1989).

Monoclonal antibodies and polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Typically, polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against non-angiogenesis regulatory proteins, using a competitive binding immunoassay. Specific polyclonal antisera and monoclonal antibodies will usually bind with a K_d of at least about 0.1 mM, more usually at least about 1 μ M, preferably at least about 0.1 μ M or better, and most preferably, 0.01 μ M or better. Antibodies specific only for a particular angiogenesis regulatory protein ortholog, such as human angiogenesis regulatory protein, can also be made, by subtracting out other cross-reacting orthologs from a species such as a non-human mammal. In this manner, antibodies that bind only to angiogenesis regulatory proteins may be obtained.

Once the specific antibodies against specific angiogenesis regulatory proteins are available, the proteins can be detected by a variety of immunoassay methods. In addition, the antibodies can be used therapeutically as a angiogenesis regulatory protein modulators. For a review of immunological and immunoassay procedures, see *Basic and Clinical Immunology* (Stites & Terr eds., 7th ed. 1991). Moreover, the immunoassays of the present

invention can be performed in any of several configurations, which are reviewed extensively in Enzyme Immunoassay (Maggio, ed., 1980); and Harlow & Lane, *supra*.

B. Immunological binding assays

5 Angiogenesis regulatory proteins can be detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.*, U.S. Patents 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see also *Methods in Cell Biology: Antibodies in Cell Biology*, volume 37 (Asai, ed. 1993); *Basic and Clinical Immunology* (Stites & Terr, eds., 7th ed. 1991). Immunological binding
10 assays (or immunoassays) typically use an antibody that specifically binds to a protein or antigen of choice (in this case the angiogenesis regulatory proteins or antigenic subsequence thereof). The antibody (e.g., anti- angiogenesis regulatory protein) may be produced by any of a number of means well known to those of skill in the art and as described above.

 Immunoassays also often use a labeling agent to specifically bind to and label
15 the complex formed by the antibody and antigen. The labeling agent may itself be one of the moieties comprising the antibody/antigen complex. Thus, the labeling agent may be a labeled angiogenesis regulatory protein or a labeled anti-angiogenesis regulatory protein antibody. Alternatively, the labeling agent may be a third moiety, such a secondary antibody, that specifically binds to the antibody/angiogenesis regulatory protein complex (a secondary
20 antibody is typically specific to antibodies of the species from which the first antibody is derived). Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G may also be used as the label agent. These proteins exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (*see, e.g.*, Kronval *et al.*, *J. Immunol.* 111:1401-1406 (1973); Akerstrom *et al.*, *J.*
25 *Immunol.* 135:2589-2542 (1985)). The labeling agent can be modified with a detectable moiety, such as biotin, to which another molecule can specifically bind, such as streptavidin. A variety of detectable moieties are well known to those skilled in the art.

 Throughout the assays, incubation and/or washing steps may be required after each combination of reagents. Incubation steps can vary from about 5 seconds to several
30 hours, optionally from about 5 minutes to about 24 hours. However, the incubation time will depend upon the assay format, antigen, volume of solution, concentrations, and the like. Usually, the assays will be carried out at ambient temperature, although they can be conducted over a range of temperatures, such as 10°C to 40°C.

Non-competitive assay formats

Immunoassays for detecting angiogenesis regulatory proteins in samples may be either competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of antigen is directly measured. In one preferred "sandwich" assay, for example, the anti-angiogenesis regulatory protein antibodies can be bound directly to a solid substrate on which they are immobilized. These immobilized antibodies then capture angiogenesis regulatory protein present in the test sample. Angiogenesis regulatory proteins thus immobilized are then bound by a labeling agent, such as a second angiogenesis regulatory protein-antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second or third antibody is typically modified with a detectable moiety, such as biotin, to which another molecule specifically binds, e.g., streptavidin, to provide a detectable moiety.

Competitive assay formats

In competitive assays, the amount of angiogenesis regulatory proteins present in the sample is measured indirectly by measuring the amount of a known, added (exogenous) angiogenesis regulatory proteins displaced (competed away) from an anti-angiogenesis regulatory protein antibody by the unknown angiogenesis regulatory proteins present in a sample. In one competitive assay, a known amount of angiogenesis regulatory proteins is added to a sample and the sample is then contacted with an antibody that specifically binds to angiogenesis regulatory proteins. The amount of exogenous angiogenesis regulatory proteins bound to the antibody is inversely proportional to the concentration of angiogenesis regulatory proteins present in the sample. In a particularly preferred embodiment, the antibody is immobilized on a solid substrate. The amount of angiogenesis regulatory proteins bound to the antibody may be determined either by measuring the amount of angiogenesis regulatory protein present in angiogenesis regulatory proteins/antibody complex, or alternatively by measuring the amount of remaining uncomplexed protein. The amount of angiogenesis regulatory proteins may be detected by providing a labeled angiogenesis regulatory protein molecule.

A hapten inhibition assay is another preferred competitive assay. In this assay the known angiogenesis regulatory proteins is immobilized on a solid substrate. A known amount of anti-angiogenesis regulatory protein antibody is added to the sample, and the

sample is then contacted with the immobilized angiogenesis regulatory protein. The amount of anti-angiogenesis regulatory protein antibody bound to the known immobilized angiogenesis regulatory protein is inversely proportional to the amount of angiogenesis regulatory proteins present in the sample. Again, the amount of immobilized antibody may be detected by detecting either the immobilized fraction of antibody or the fraction of the antibody that remains in solution. Detection may be direct where the antibody is labeled or indirect by the subsequent addition of a labeled moiety that specifically binds to the antibody as described above.

Cross-reactivity determinations

Immunoassays in the competitive binding format can also be used for crossreactivity determinations. For example, angiogenesis regulatory proteins can be immobilized to a solid support. Proteins (e.g., angiogenesis regulatory protein and homologs) are added to the assay that compete for binding of the antisera to the immobilized antigen. The ability of the added proteins to compete for binding of the antisera to the immobilized protein is compared to the ability of the angiogenesis regulatory protein to compete with itself. The percent crossreactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% crossreactivity with each of the added proteins listed above are selected and pooled. The cross-reacting antibodies are optionally removed from the pooled antisera by immunoabsorption with the added considered proteins, e.g., distantly related homologs.

The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein, thought to be perhaps an allele or polymorphic variant of a angiogenesis regulatory proteins, to the immunogen protein. In order to make this comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. If the amount of the second protein required to inhibit 50% of binding is less than 10 times the amount of the angiogenesis regulatory proteins that is required to inhibit 50% of binding, then the second protein is said to specifically bind to the polyclonal antibodies generated to angiogenesis regulatory protein immunogen.

Other assay formats

Western blot (immunoblot) analysis is used to detect and quantify the presence of an angiogenesis regulatory protein in the sample. The technique generally comprises separating sample proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support, (such as a nitrocellulose filter, a nylon filter, or derivatized nylon filter), and incubating the sample with the antibodies that specifically bind angiogenesis regulatory protein. The anti-angiogenesis regulatory protein antibodies specifically bind to the angiogenesis regulatory protein on the solid support. These antibodies may be directly labeled or alternatively may be subsequently detected using labeled antibodies (e.g., labeled sheep anti-mouse antibodies) that specifically bind to the anti-angiogenesis regulatory protein antibodies.

Other assay formats include liposome immunoassays (LIA), which use liposomes designed to bind specific molecules (e.g., antibodies) and release encapsulated reagents or markers. The released chemicals are then detected according to standard techniques (see Monroe *et al.*, *Amer. Clin. Prod. Rev.* 5:34-41 (1986)).

Reduction of non-specific binding

One of skill in the art will appreciate that it is often desirable to minimize non-specific binding in immunoassays. Particularly, where the assay involves an antigen or antibody immobilized on a solid substrate it is desirable to minimize the amount of non-specific binding to the substrate. Means of reducing such non-specific binding are well known to those of skill in the art. Typically, this technique involves coating the substrate with a proteinaceous composition. In particular, protein compositions such as bovine serum albumin (BSA), nonfat powdered milk, and gelatin are widely used with powdered milk being most preferred.

Labels

The particular label or detectable group used in the assay is not a critical aspect of the invention, as long as it does not significantly interfere with the specific binding of the antibody used in the assay. The detectable group can be any material having a detectable physical or chemical property. Such detectable labels have been well-developed in the field of immunoassays and, in general, most any label useful in such methods can be applied to the present invention. Thus, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical

means. Useful labels in the present invention include magnetic beads (e.g., DYNABEADSTM), fluorescent dyes (e.g., fluorescein isothiocyanate, Texas red, rhodamine, and the like), radiolabels (e.g., ³H, ¹²⁵I, ³⁵S, ¹⁴C, or ³²P), enzymes (e.g., horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric
5 labels such as colloidal gold or colored glass or plastic beads (e.g., polystyrene, polypropylene, latex, etc.).

The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. As indicated above, a wide variety of labels may be used, with the choice of label depending on sensitivity required, ease of
10 conjugation with the compound, stability requirements, available instrumentation, and disposal provisions.

Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to another molecules (e.g., streptavidin) molecule, which is either inherently detectable or
15 covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. The ligands and their targets can be used in any suitable combination with antibodies that recognize angiogenesis regulatory proteins, or secondary antibodies that recognize anti-angiogenesis regulatory protein.

The molecules can also be conjugated directly to signal generating
20 compounds, e.g., by conjugation with an enzyme or fluorophore. Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidotases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol. For a review
25 of various labeling or signal producing systems that may be used, see U.S. Patent No. 4,391,904.

Means of detecting labels are well known to those of skill in the art. Thus, for example, where the label is a radioactive label, means for detection include a scintillation counter or photographic film as in autoradiography. Where the label is a fluorescent label, it
30 may be detected by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence. The fluorescence may be detected visually, or by the use of electronic detectors such as charge coupled devices (CCDs) or photomultipliers and the like. Similarly, enzymatic labels may be detected by providing the appropriate substrates for the enzyme and detecting the resulting reaction product. Finally simple colorimetric labels

may be detected simply by observing the color associated with the label. Thus, in various dipstick assays, conjugated gold often appears pink, while various conjugated beads appear the color of the bead.

Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labeled and the presence of the target antibody is detected by simple visual inspection.

10 Gene therapy

The present invention provides the nucleic acids of angiogenesis regulatory protein for the transfection of cells *in vitro* and *in vivo*. These nucleic acids can be inserted into any of a number of well-known vectors for the transfection of target cells and organisms as described below. The nucleic acids are transfected into cells, *ex vivo* or *in vivo*, through the interaction of the vector and the target cell. The nucleic acid, under the control of a promoter, then expresses the angiogenesis regulatory proteins of the present invention, thereby mitigating the effects of absent, partial inactivation, or abnormal expression of the angiogenesis regulatory protein gene, particularly as it relates to angiogenesis. The compositions are administered to a patient in an amount sufficient to elicit a therapeutic response in the patient. An amount adequate to accomplish this is defined as "therapeutically effective dose or amount."

Such gene therapy procedures have been used to correct acquired and inherited genetic defects, cancer, and other diseases in a number of contexts. The ability to express artificial genes in humans facilitates the prevention and/or cure of many important human diseases, including many diseases which are not amenable to treatment by other therapies (for a review of gene therapy procedures, *see* Anderson, *Science* 256:808-813 (1992); Nabel & Felgner, *TIBTECH* 11:211-217 (1993); Mitani & Caskey, *TIBTECH* 11:162-166 (1993); Mulligan, *Science* 926-932 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10):1149-1154 (1998); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer & Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada *et al.*, in *Current Topics in Microbiology and Immunology* (Doerfler & Böhm eds., 1995); and Yu *et al.*, *Gene Therapy* 1:13-26 (1994)).

The nucleic acids of the invention can also be used to make transgenic animals, such as transgenic mice, either by knock-out or overexpression. Such animals are useful, e.g., for testing modulators of angiogenesis.

5 **Pharmaceutical compositions and administration**

Pharmaceutically acceptable carriers are determined in part by the particular composition being administered (e.g., nucleic acid, protein, modulatory compounds or transduced cell), as well as by the particular method used to administer the composition. Accordingly, there are a wide variety of suitable formulations of pharmaceutical compositions of the present invention (*see, e.g., Remington's Pharmaceutical Sciences*, 17th ed., 1989). Administration can be in any convenient manner, e.g., by injection, oral administration, inhalation, transdermal application, or rectal administration.

Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of the packaged nucleic acid suspended in diluents, such as water, saline or PEG 400; (b) capsules, sachets or tablets, each containing a predetermined amount of the active ingredient, as liquids, solids, granules or gelatin; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. Tablet forms can include one or more of lactose, sucrose, mannitol, sorbitol, calcium phosphates, corn starch, potato starch, microcrystalline cellulose, gelatin, colloidal silicon dioxide, talc, magnesium stearate, stearic acid, and other excipients, colorants, fillers, binders, diluents, buffering agents, moistening agents, preservatives, flavoring agents, dyes, disintegrating agents, and pharmaceutically compatible carriers. Lozenge forms can comprise the active ingredient in a flavor, e.g., sucrose, as well as pastilles comprising the active ingredient in an inert base, such as gelatin and glycerin or sucrose and acacia emulsions, gels, and the like containing, in addition to the active ingredient, carriers known in the art.

The compound of choice, alone or in combination with other suitable components, can be made into aerosol formulations (i.e., they can be "nebulized") to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

Formulations suitable for parenteral administration, such as, for example, by intraarticular (in the joints), intravenous, intramuscular, intradermal, intraperitoneal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile

suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. In the practice of this invention, compositions can be administered, for example, by intravenous infusion, orally, topically, intraperitoneally, intravesically or intrathecally. Parenteral administration and intravenous administration are the preferred methods of administration. The formulations of commends can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials.

Injection solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described. Cells transduced by nucleic acids for *ex vivo* therapy can also be administered intravenously or parenterally as described above.

The dose administered to a patient, in the context of the present invention should be sufficient to effect a beneficial therapeutic response in the patient over time. The dose will be determined by the efficacy of the particular vector employed and the condition of the patient, as well as the body weight or surface area of the patient to be treated. The size of the dose also will be determined by the existence, nature, and extent of any adverse side-effects that accompany the administration of a particular vector, or transduced cell type in a particular patient.

In determining the effective amount of the vector to be administered in the treatment or prophylaxis of conditions owing to diminished or aberrant expression of the angiogenesis regulatory proteins, the physician evaluates circulating plasma levels of the vector, vector toxicities, progression of the disease, and the production of anti-vector antibodies. In general, the dose equivalent of a naked nucleic acid from a vector is from about 1 μg to 100 μg for a typical 70 kilogram patient, and doses of vectors which include a retroviral particle are calculated to yield an equivalent amount of therapeutic nucleic acid.

For administration, compounds and transduced cells of the present invention can be administered at a rate determined by the LD-50 of the inhibitor, vector, or transduced cell type, and the side-effects of the inhibitor, vector or cell type at various concentrations, as applied to the mass and overall health of the patient. Administration can be accomplished via single or divided doses.

EXAMPLES

Using a functional assay described herein (see also Figure 1), angiogenesis regulatory proteins and nucleic acids have been identified and cloned from a green fluorescent protein (GFP) fusion cDNA retroviral library. The cDNAs can be in-frame, out of frame, sense, or antisense. Primary endothelial cells were transduced with retroviral

cDNA/GFP fusion libraries. Transduced cells were selected and assayed for alterations consistent with modulation of angiogenesis, *e.g.* downregulation of cell surface expression of $\alpha v \beta 3$ or downregulation of haptotaxis.

5 Example 1: Identification of plakoglobin angiogenesis regulatory protein.

Using the functional assay described above, expression of a plakoglobin-GFP fusion protein (*i.e.*, GFP- Δ N-plakoglobin) was found to decrease levels of the cell surface marker $\alpha v \beta 3$. (*See, e.g.*, Figures 2 and 3.) The plakoglobin functional hit was an N-terminal deletion, fused in frame to the GFP protein that contained armadillo repeats. (*See, e.g.*,
10 Figure 3). GFP- Δ N-Plakoglobin expression downregulated vitronectin receptors ($\alpha v \beta 3$ and $\alpha v \beta 5$) but not the fibronectin receptor ($\alpha 5 \beta 1$). (*See, e.g.*, Figure 4.)

The angiogenesis phenotype of GFP- Δ N-Plakoglobin was confirmed by co-culturing HUVECs on smooth muscle cells and using a tubule formation assay. Briefly, on day 1 a Collagen solution (Collagen type I solution, Sigma C8919) was diluted 1 to 10 with
15 PBS and used to coat each well of 24 well plate (250ul). Plates were incubated at 37°C for 4-5 hrs. Wells were then aspirated and allowed to dry overnight. On day 2, stock flask(s) of PSMCs (Clonetics Pulmonary Artery SMCs CC-2581, fewer than 5 passages) were trypsinized to provide enough cells to plate 60,000 cells per well. Cells were cultured in Smooth muscle cell growth medium (Clonetics SmGM-2 bullet kit CC-3182) and stock
20 cultures were maintained according to Clonetics protocols.

PSMCs were nearly confluent on day 3. Trypsinize Stock flask(s) of HUVECs were trypsinized to generate enough cells to plate 12,000 cells per well (trypsin, Clonetics ReagentPack CC-5034). The SmGM-2 medium was removed and HUVECS were plated onto the PSMCs in EGM-2 (Endothelial cell growth medium, Clonetics EGM-2
25 bullet kit CC-3162) at 1ml per well.

Medium was exchanged every 2-3 days and tubule formation was assessed. Using the tubule formation assay, GFP- Δ N-Plakoglobin expression was found to inhibit branch formation. (Data not shown.)

Expression of GFP- Δ N-Plakoglobin downregulated endogenous β -catenin
30 levels, as well as expression of endogenous plakoglobin and VE-cadherin. (*See, e.g.*, Figure 5.) β -catenin was shown to be required to maintain $\alpha v \beta 3$ expression. (Data not shown.)

In addition, GFP- Δ N-Plakoglobin expression upregulated surface levels of the laminin receptor, $\alpha 6 \beta 4$. (*See, e.g.*, Figure 6.)

Example 2: Identification of HoxB2 angiogenesis regulatory protein.

Using the functional assay described above, expression of a HoxB2-GFP fusion protein (*i.e.*, Δ N-HoxB2) was found to decrease levels of the cell surface marker α v β 3. Cells transfected with the Δ N-HoxB2-GFP fusion protein construct exhibit lower α v β 3 expression levels than control cells. The HoxB2 functional screen hit is an N-terminal deletion that contains the HoxB2 homeodomain. (*See, e.g.*, Figure 7.)

Δ N-HoxB2 downregulated both α v and β 3 integrin subunit surface expression, as well as α v β 5 surface expression. (*See, e.g.*, Figure 8.) Δ N-HoxB2 also downregulated β 3 transcript expression. (Data not shown.) Δ N-HoxB2 inhibited proliferation in HUVECs, a primary cell line. (*See, e.g.*, Figure 9.) In addition, full length HoxB2 phenocopied Δ N-HoxB2 and downregulated surface expression of α v β 3. (*See, e.g.*, Figure 10.)

Cells expressing full length HoxB2 or Δ N-HoxB2 were isolated by cell sorting and assayed for tube formation using the co-culture assay described above. Both full length HoxB2 and Δ N-HoxB2 increased branch formation in the co-culture assay. (Data not shown.) α v β 3 expression was directly downregulated using β 3 siRNA. Downregulation of α v β 3 expression did not affect branch formation in the co-culture assay.

Example 3: Identification of SUSP-1 angiogenesis regulatory protein.

Using the functional assay described above, expression of a SUSP-1 GFP fusion protein (*i.e.*, GFP-SUSP-1) was found to decrease levels of the cell surface marker α v β 3. (*See, e.g.*, Figure 11.) Cells transfected with the GFP-SUSP-1 fusion protein construct exhibit lower α v β 3 expression levels than control cells. The SUSP-1 functional screen hit is an N-terminal fusion in the sense orientation and that has an additional encoded cDNA insert before the wild-type start codon. (*See, e.g.*, Figure 11.)

The GFP-SUSP-1 screening hit did not affect proliferation when expressed in HUVEC (human umbilical vein endothelial cells), PASMCM (pulmonary artery smooth muscle cells), and NHDF (normal human dermal fibroblasts). (*See, e.g.*, Figure 12.) SUSP-1 mRNA was found to be expressed in a variety of cell lines. (*See, e.g.*, Figure 13.)

GFP-SUSP-1 expression inhibited tube formation in the co-culture assay described above. (Data not shown.) Two SUSP-1 siRNAs were synthesized and found to inhibit SUSP-1 expression. (*See, e.g.*, Figure 14.) The SUSP-1 siRNAs also inhibited tube formation in the co-culture assay. (Data not shown.)

Example 4: Identification of $\Delta 5$ FADS angiogenesis regulatory protein.

Using the functional assay described above, expression of a $\Delta 5$ FADS antisense nucleic acid (*i.e.*, GFP- $\Delta 5$ FADS) was found to decrease levels of the cell surface marker $\alpha v \beta 3$. (*See, e.g.*, Figure 30.) Cells transfected with the GFP- $\Delta 5$ FADS construct exhibit lower $\alpha v \beta 3$ expression levels than control cells. The $\Delta 5$ FADS functional screen hit is a 300bp antisense nucleic acid. (*See, e.g.*, Figure 30.)

The GFP- $\Delta 5$ FADS screening hit did not affect proliferation when expressed in HUVEC (human umbilical vein endothelial cells), PASMC (pulmonary artery smooth muscle cells), and NHDF (normal human dermal fibroblasts). (*See, e.g.*, Figure 31.) $\Delta 5$ FADS mRNA was found to be expressed in a variety of cell lines, and appeared to be highly expressed in endothelial cells, hepatocytes, and mammary epithelial cells. (Data not shown.)

Antisense $\Delta 5$ FADS nucleic acids inhibited tube formation in the co-culture assay described above. (Data not shown.) Two $\Delta 5$ FADS siRNAs were synthesized and found to inhibit $\Delta 5$ FADS expression. (*See, e.g.*, Figure 32.) The $\Delta 5$ FADS siRNAs inhibited tube formation in the co-culture assay. (Data not shown.) The $\Delta 5$ FADS siRNAs inhibited vitronectin haptotaxis and reduced SDF-1 induced chemotaxis in Jurkat cells. (*See, e.g.*, Figures 33-34.)

Example 5: Identification of $\Delta 6$ FADS angiogenesis regulatory protein.

Using the functional assay described above, expression of a KIAA1583 antisense nucleic acid (*i.e.*, GFP-KIAA1583) was found to decrease levels of the cell surface marker $\alpha v \beta 3$. (Data not shown.) KIAA1583 and $\Delta 6$ FADS genomic sequences are both located on chromosome 11q12, and transcripts of the two genes are spiced together. (*See, e.g.*, Figure 35.) $\Delta 6$ FADS mRNA is reduced in cells expressing GFP-KIAA1583, the antisense screening hit. (*See, e.g.*, Figure 36.) $\Delta 6$ FADS also interacts with C1 phosphatase, another $\alpha v \beta 3$ screening hit. (Data not shown.) $\Delta 6$ FADS is highly expressed in endothelial cells, hepatocytes and skeletal muscle cells. (Data not shown.)

$\Delta 6$ FADS siRNAs were synthesized and found to inhibit $\Delta 5$ FADS expression. (*See, e.g.*, Figure 37.) The $\Delta 6$ FADS siRNAs inhibited tube formation in the co-culture assay described above. (Data not shown.) Expression of the $\Delta 6$ FADS siRNAs also downregulated $\alpha v \beta 3$ surface expression levels. (*See, e.g.*, Figure 37.)

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

5 Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to one of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

WHAT IS CLAIMED IS:

- 1 1. A method for identifying a compound that modulates angiogenesis,
2 the method comprising the steps of:
- 3 (i) contacting the compound with a nucleic acid, or a polypeptide or a
4 fragment thereof encoded by a nucleic acid, wherein the nucleic acid hybridizes under
5 stringent conditions to a second nucleic acid comprising a nucleotide sequence selected
6 from the group consisting of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID
7 NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID
8 NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID
9 NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID
10 NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID
11 NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID
12 NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID
13 NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID
14 NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID
15 NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID
16 NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID
17 NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID
18 NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ
19 ID NO:523 ; and
- 20 (ii) determining the functional effect of the compound upon the nucleic
21 acid or polypeptide.
- 1 2. The method of claim 1, wherein the functional effect is determined
2 *in vitro*.
- 1 3. The method of claim 2, wherein the functional effect is a physical
2 effect.
- 1 4. The method of claim 2, wherein the functional effect is determined
2 by measuring ligand or substrate binding to the polypeptide.
- 1 5. The method of claim 2, wherein the functional effect is a chemical
2 effect.

1 6. The method of claim 2, wherein the functional effect is determined
2 by measuring an enzymatic activity.

1 7. The method of claim 1, wherein the polypeptide is expressed in a
2 eukaryotic host cell.

1 8. The method of claim 7, wherein the functional effect is a physical
2 effect.

1 9. The method of claim 8, wherein the functional effect is determined
2 by measuring ligand or substrate binding to the polypeptide.

1 10. The method of claim 7, wherein the functional effect is a chemical
2 or phenotypic effect.

1 11. The method of claim 10, wherein the functional effect is
2 determined by measuring an enzymatic activity.

1 12. The method of claim 7, wherein the host cell is an endothelial cell.

1 13. The method of claim 12, wherein the functional effect is
2 determined by measuring $\alpha v\beta 3$ expression.

1 14. The method of claim 12, wherein the functional effect is
2 determined by measuring a member selected from $\alpha v\beta 3$ expression, haptotaxis, and
3 chemotaxis.

1 15. The method of claim 14, wherein the functional effect is
2 determined by measuring $\alpha v\beta 3$ expression.

1 16. The method of claim 1, wherein modulation is inhibition of
2 angiogenesis.

1 17. The method of claim 1 wherein the polypeptide is recombinant.

1 18. The method of claim 1, wherein the nucleic acid comprises a
2 nucleotide sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:32,
3 SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ

4 ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID
5 NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID
6 NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID
7 NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID
8 NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID
9 NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID
10 NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID
11 NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID
12 NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID
13 NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID
14 NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID
15 NO:521, and SEQ ID NO:523.

1 19. The method of claim 1, wherein the polypeptide comprises a
2 sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:33, SEQ ID
3 NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID
4 NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID
5 NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID
6 NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID
7 NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID
8 NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID
9 NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID
10 NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID
11 NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID
12 NO:520, SEQ ID NO:522, and SEQ ID NO:524.

1 20. The method of claim 1, wherein the compound is an antibody.

1 21. The method of claim 1, wherein the compound is an antisense
2 molecule.

1 22. The method of claim 1, wherein the compound is a small organic
2 molecule.

1 23. The method of claim 1, wherein the compound is a peptide.

1 24. A method for identifying a compound that modulates angiogenesis,
2 the method comprising the steps of:

3 (i) contacting the compound with a nucleic acid, or a polypeptide or a
4 fragment thereof encoded by a nucleic acid, wherein the nucleic acid hybridizes under
5 stringent conditions to a second nucleic acid comprising a nucleotide sequence selected
6 from the group consisting of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID
7 NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID
8 NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID
9 NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID
10 NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID
11 NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID
12 NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID
13 NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID
14 NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID
15 NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID
16 NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439 SEQ ID
17 NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID
18 NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ
19 ID NO:523;

20 (ii) determining the functional effect of the compound upon the nucleic
21 acid or polypeptide; and

22 (iii) expressing the nucleic acid or polypeptide in a cell, contacting the
23 nucleic acid or polypeptide with the compound, and determining the phenotypic or
24 chemical effect upon the cell.

1 25. A method of modulating angiogenesis in a subject, the method
2 comprising the step of administering to the subject a therapeutically effective amount of a
3 compound identified using the method of claim 1.

1 26. The method of claim 25, wherein the subject is a human.

1 27. The method of claim 25, wherein the compound is an antibody.

1 28. The method of claim 25, wherein the compound is an antisense
2 molecule.

1 29. The method of claim 25, wherein the compound is a small organic
2 molecule.

1 30. The method of claim 25, wherein the compound is a peptide.

1 31. The method of claim 25, wherein the compound is an RNAi
2 molecule.

1 32. The method of claim 25, wherein the compound inhibits
2 angiogenesis

1 33. A method of modulating angiogenesis in a subject, the method
2 comprising the step of administering to the subject a therapeutically effective amount of a
3 polypeptide, the polypeptide encoded by a nucleic acid that hybridizes under stringent
4 conditions to a second nucleic acid comprising a nucleotide sequence selected from the
5 group consisting of SEQ ID NO:63, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86,
6 SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:165, SEQ ID NO:183,
7 SEQ ID NO:202, SEQ ID NO:218, SEQ ID NO:232, SEQ ID NO:274, SEQ ID NO:285,
8 SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320,
9 SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:340, SEQ ID NO:377, SEQ ID NO:384,
10 SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:439 SEQ ID NO:445, SEQ ID NO:456,
11 SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, and SEQ ID
12 NO:498.

1 34. A method of modulating angiogenesis in a subject, the method
2 comprising the step of administering to the subject a therapeutically effective amount of a
3 nucleic acid that hybridizes under stringent conditions to a second nucleic acid
4 comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:3,
5 SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ
6 ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID
7 NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID
8 NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID
9 NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID
10 NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID
11 NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID

12 NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID
13 NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID
14 NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID
15 NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID
16 NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID
17 NO:519, SEQ ID NO:521, and SEQ ID NO:523.

1 35. An isolated nucleic acid, wherein the nucleic acid hybridizes
2 under stringent conditions to a polynucleotide comprising a nucleotide sequence selected
3 from the group consisting of SEQ ID NO:165, SEQ ID NO:202, SEQ ID NO:210, SEQ
4 ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ
5 ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ
6 ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ
7 ID:329, and SEQ ID:330.

1 36. The isolated nucleic acid of claim 35, wherein the nucleic acid
2 comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:165,
3 SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232,
4 SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297,
5 SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320,
6 SEQ ID NO:323, SEQ ID NO:324, SEQ ID:329, and SEQ ID:330.

1 37. An isolated polypeptide comprising an amino acid sequence
2 selected from the group consisting of SEQ ID NO:287, SEQ ID NO:298, SEQ ID
3 NO:309, SEQ ID NO:319, SEQ ID NO:325, and SEQ ID NO:331.

THIS PAGE BLANK (USPTO)

Anti-Angiogenesis Screening: Target Angiogenic-Endothelial Cell Functions

WO 2004/019893

10/525573

PCT/US2003/027523

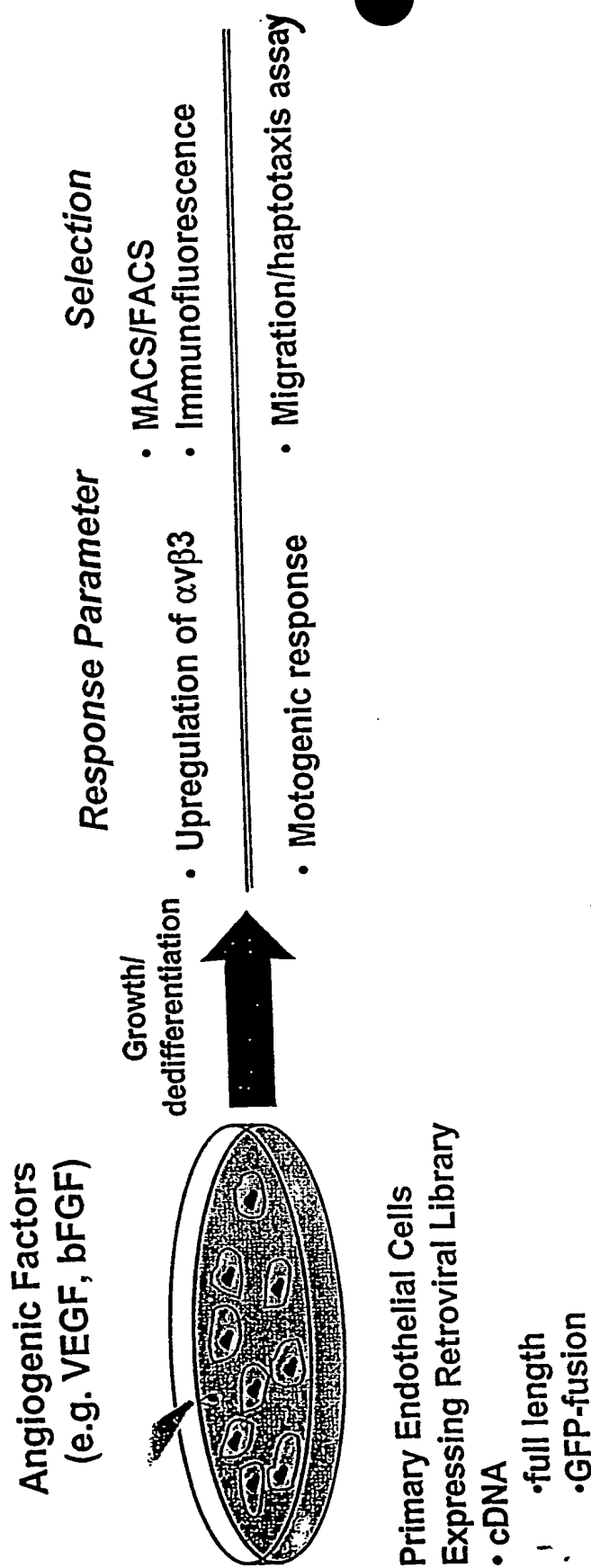
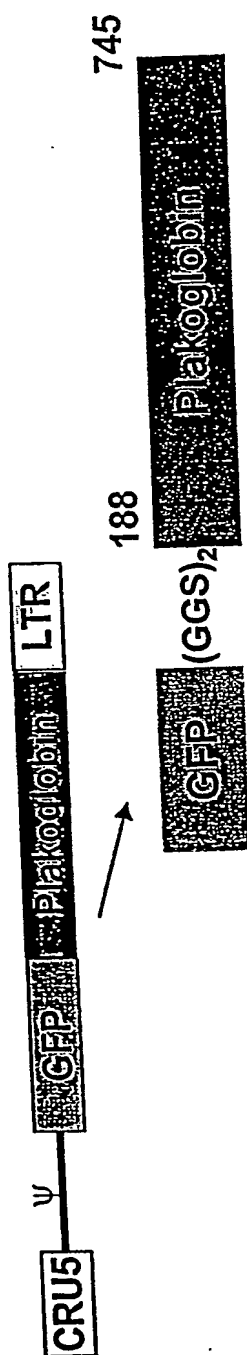


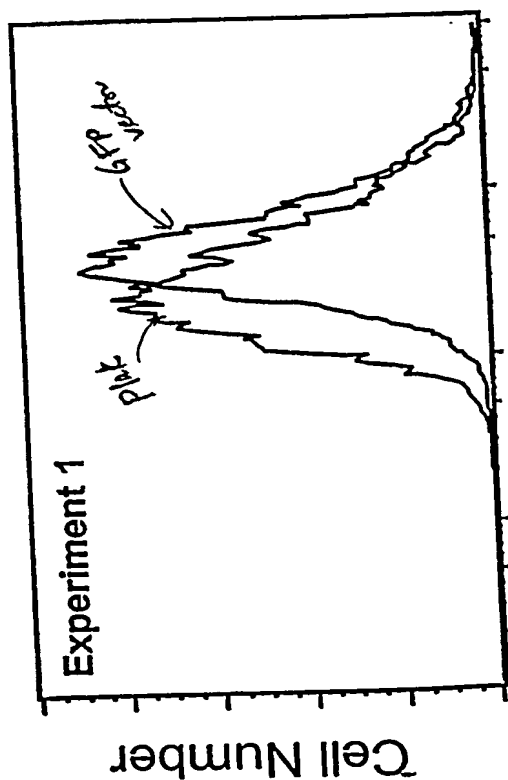
FIG. 1

THIS PAGE BLANK (USPTO)

Clone 19B5 Encodes a GFP-Plakoglobin Fusion Protein



— GFP-Plakoglobin
— GFP-vector



$\alpha v \beta 3$

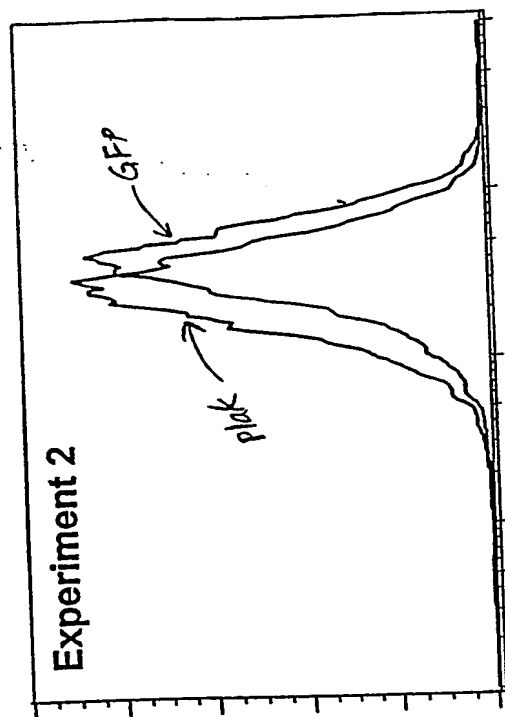


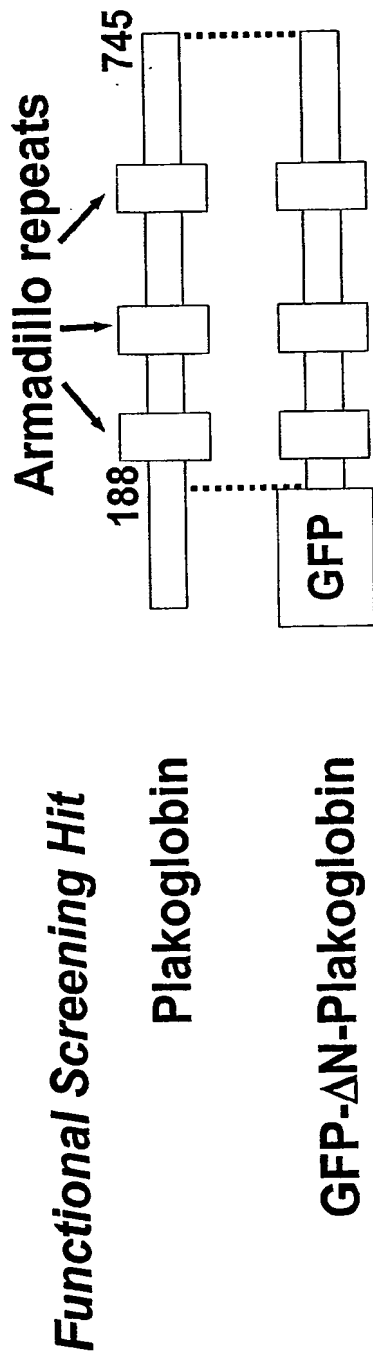
FIG. 2

2 10 1995 11 23 FEB 2005

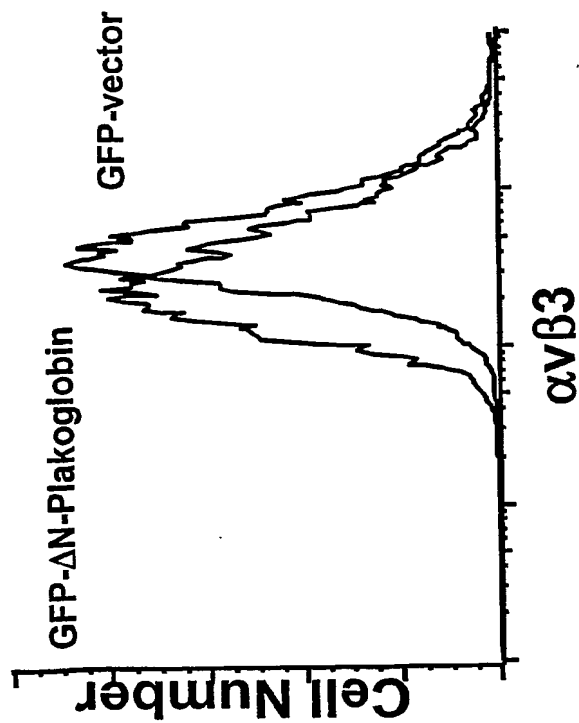
THIS PAGE BLANK (USPTO)

GFP-ΔN-Plakoglobin Expression in Endothelial Cells Downregulates α v β 3 Surface Levels

Figure 3



FACS Analysis



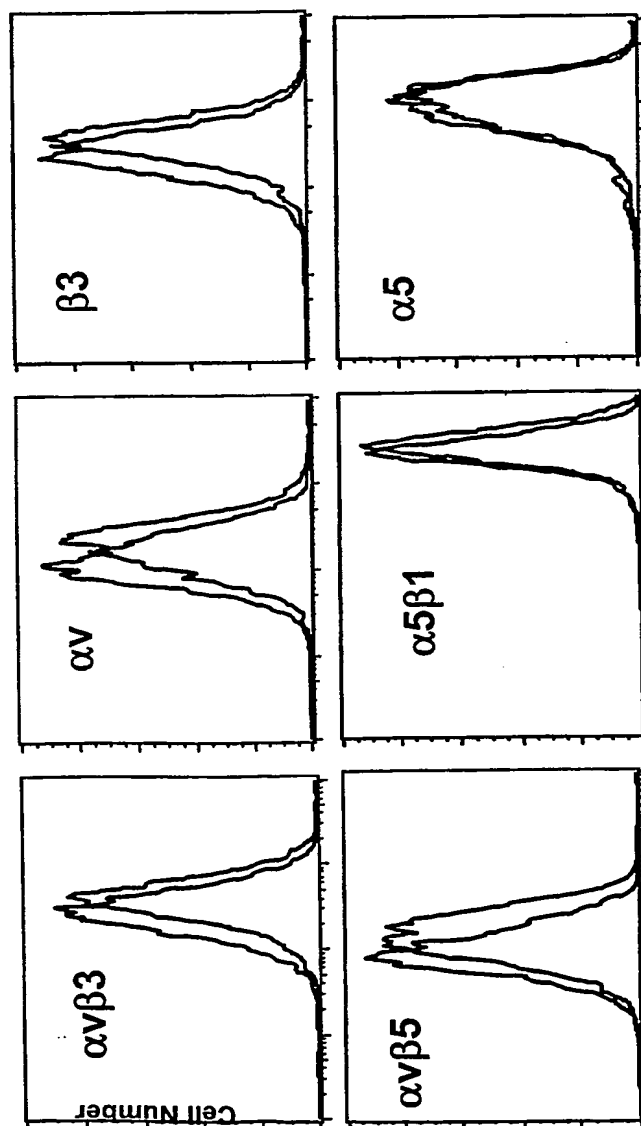
DT15 Rec 23 FEB 2005

THIS PAGE BLANK (USPTO)

Figure 4

GFP-ΔN-Plako Expression Downregulates Vitronectin Receptors (αvβ3 and αvβ5) But Not the Fibronectin Receptor (α5β1)

FACS Analysis



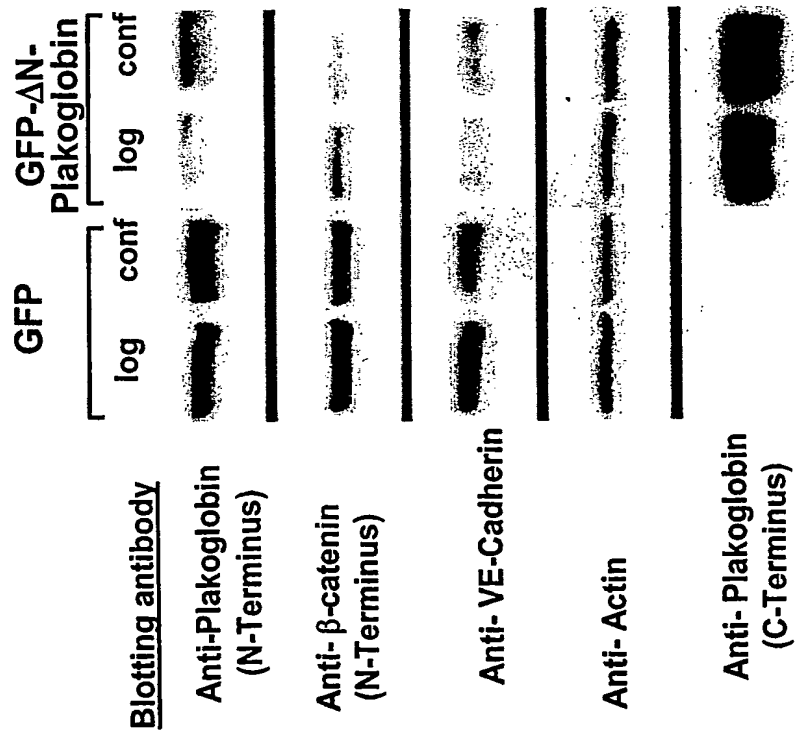
DT15 Rec'd USPTO 23 FEB 2005

THIS PAGE BLANK (USPTO)

Figure 5

Expression of GFP- Δ N-Plakoglobin Downregulates Endogenous Catenin Levels

Western blot

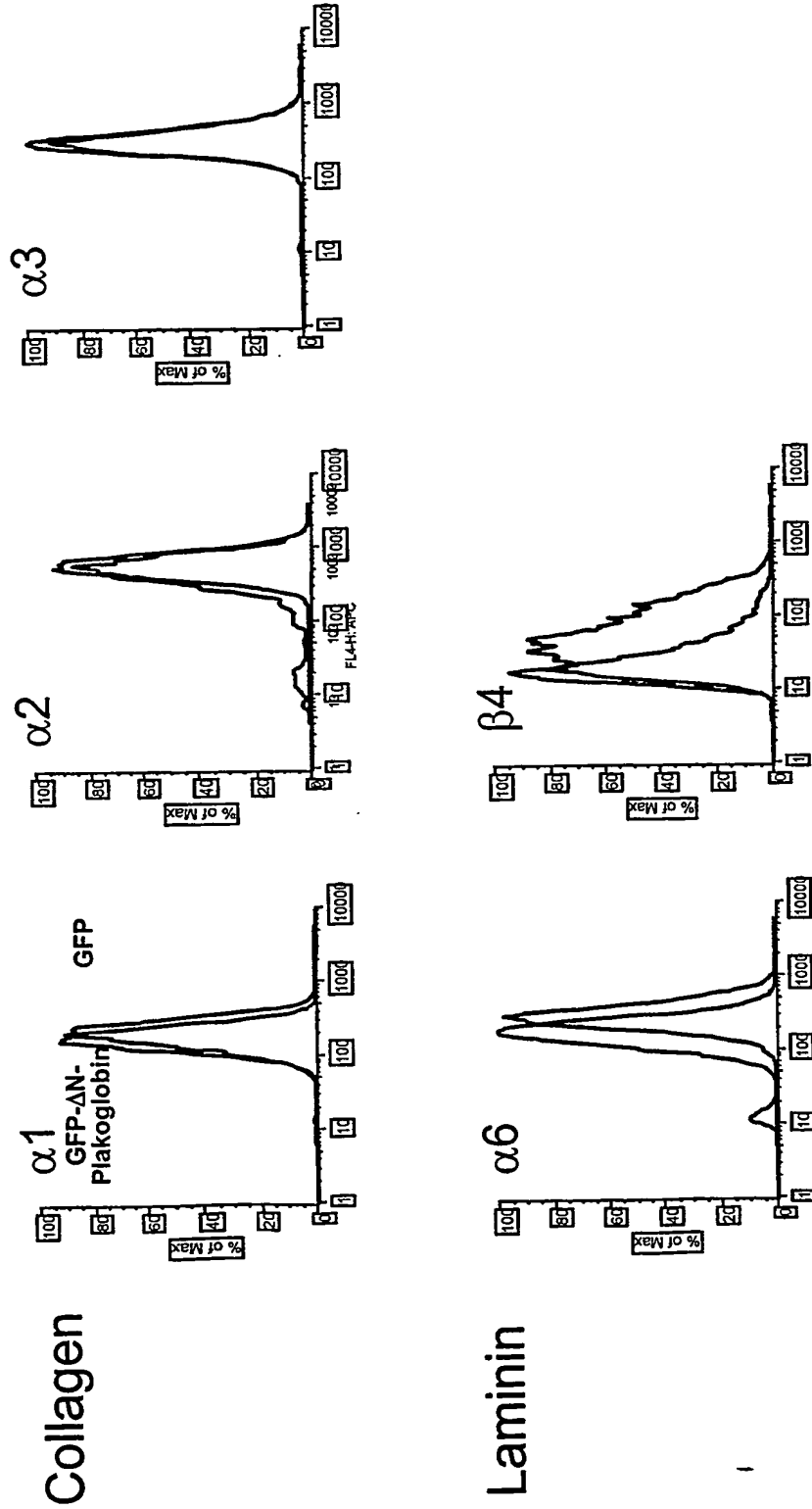


(Transduced HUVEC)

THIS PAGE BLANK (USPTO)

GFP-ΔN-Plakoglobin Expression Upregulates Surface Levels of the Laminin Receptor, α6β4

Figure 6



23 JUL 2005

THIS PAGE BLANK (USPTO)

αvβ3 Screen Clone Encodes an N-terminally Truncated Splice Variant of HoxB2

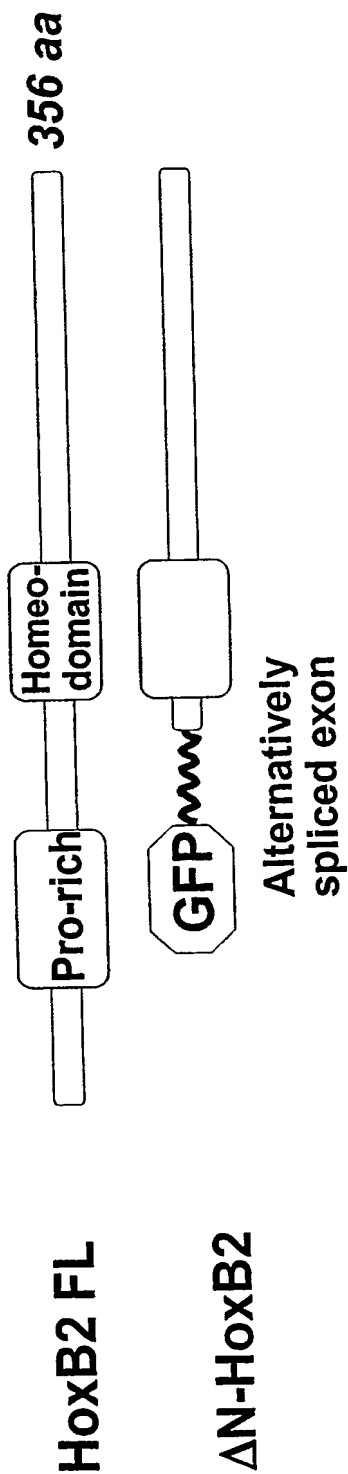
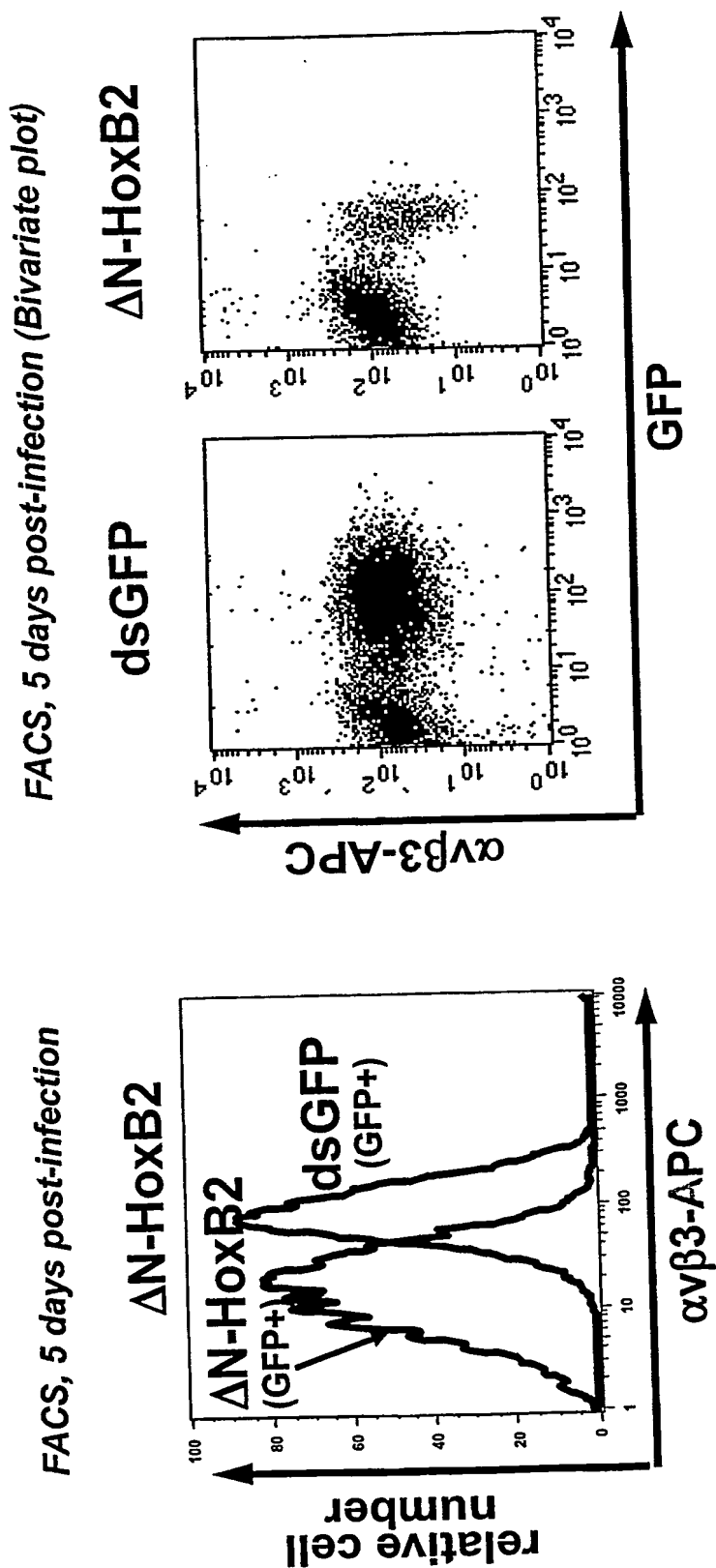


Figure 7



23 FEB 2005

THIS PAGE IS BLANK

Δ N-HoxB2 Downregulates Both α v and β 3 Integrin Subunit Surface Expression

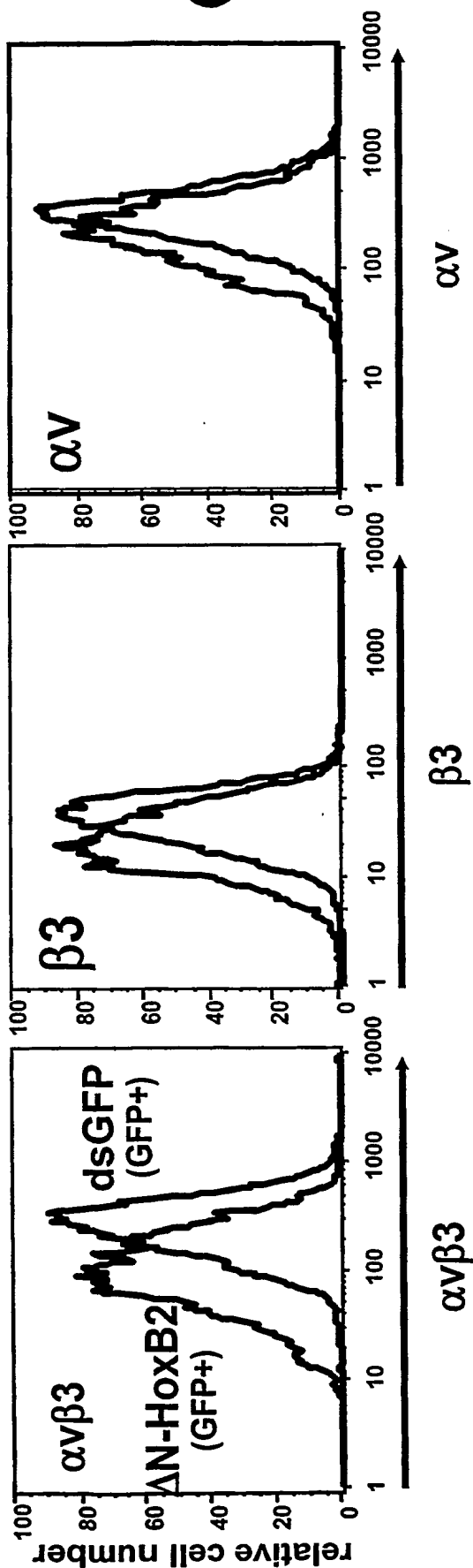
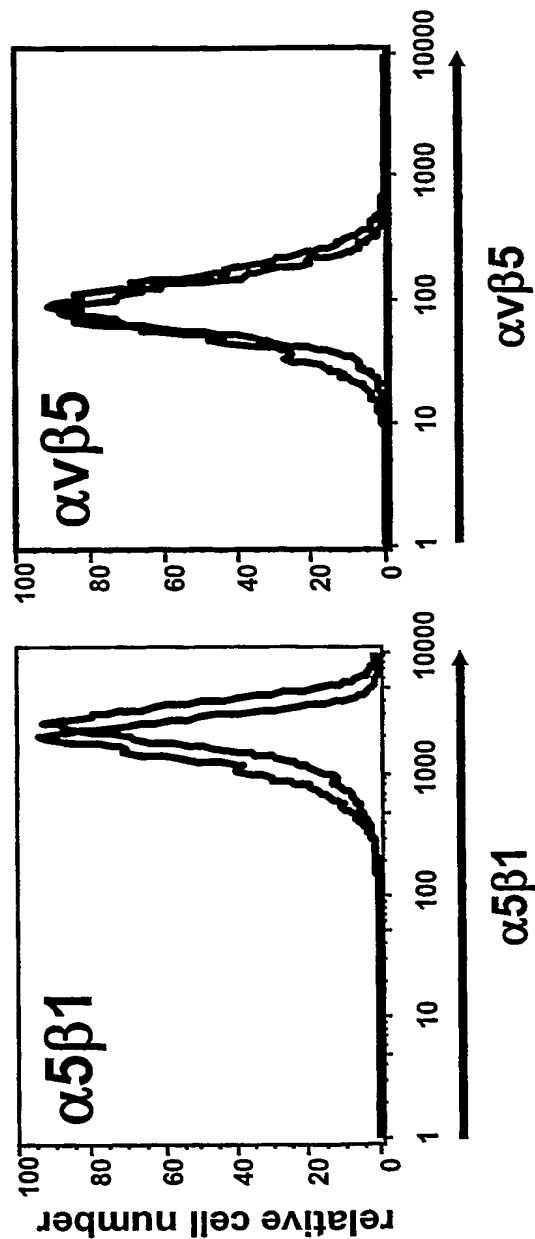


Figure 8

Δ N-HoxB2 also Downregulates α 5 β 1 Surface Expression



DT15 Rep DTIC 23 FEB 2005

THIS PAGE RI ANK DTIC

Δ N-HoxB2 Inhibits Proliferation in HUVECs

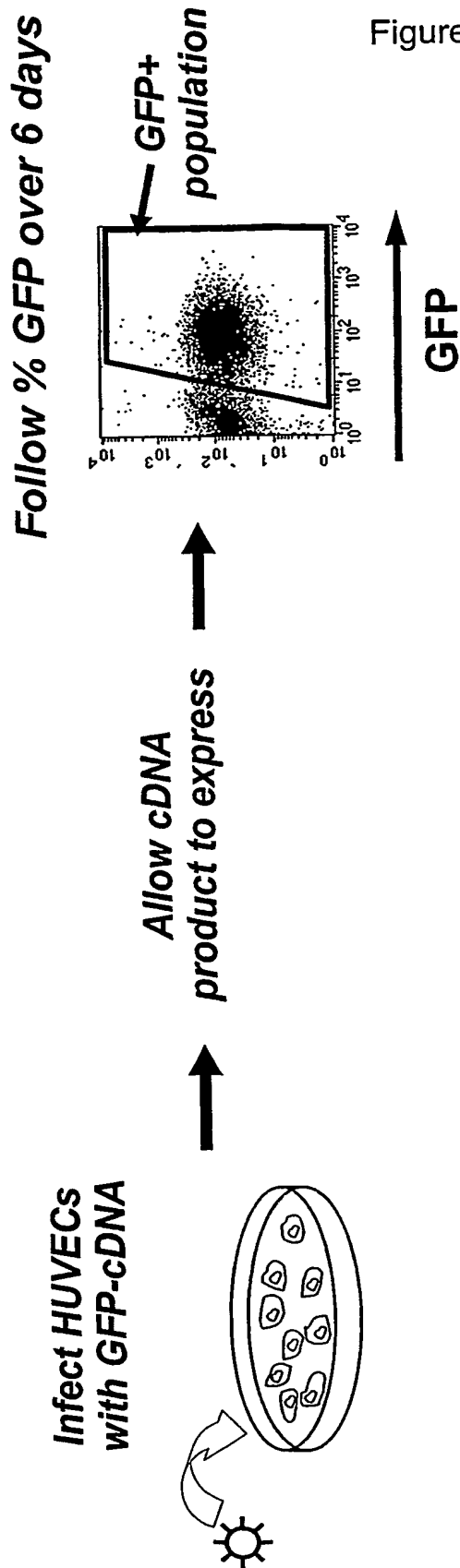
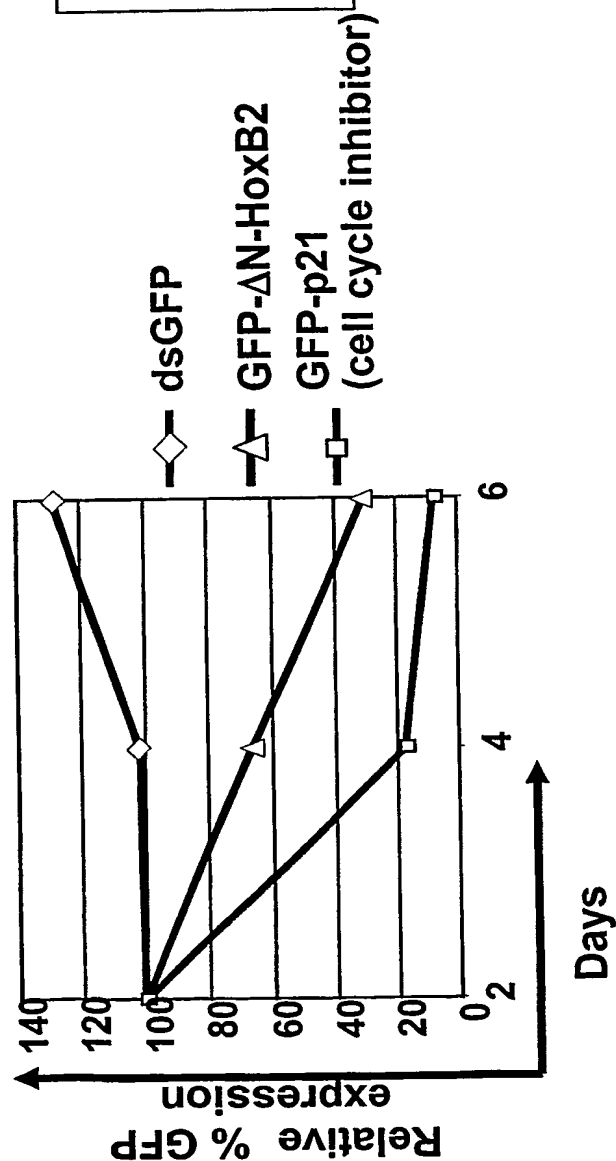


Figure 9



% GFP+ Δ N-HoxB2-expressing cells decreases over time compared to dsGFP vector, indicative of decreased proliferation (or increased apoptosis)

DT15 Rec'd PCT/PTO 23 FEB 2005

THIS PAGE BLANK (USPTO)

Full Length HoxB2 Phenocopies Δ N-HoxB2 in the Downregulation of α v β 3

WO 2004/019893

10/525573

PCT/US2003/027523

FACS, 5 days post-infection

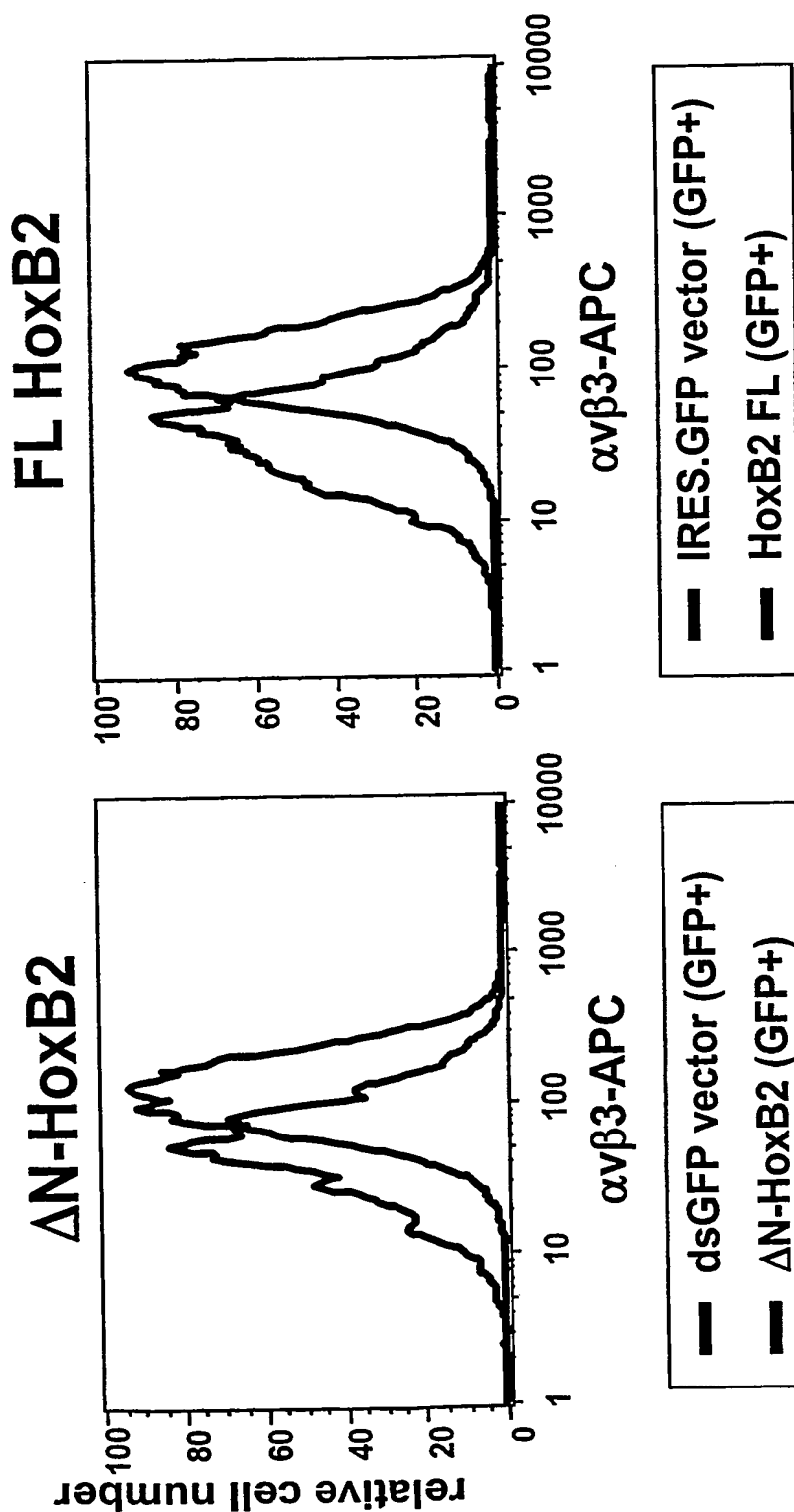


Figure 10

23 SEP 2008

THIS PAGE BLANK (ISPTO)

SUSP-1

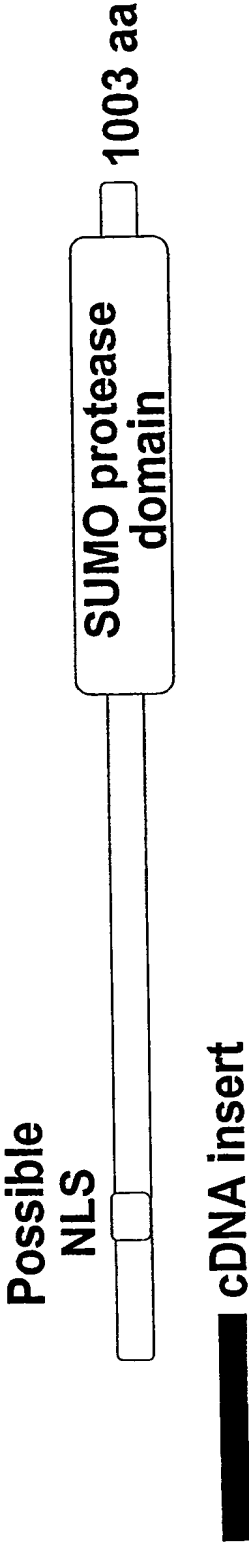
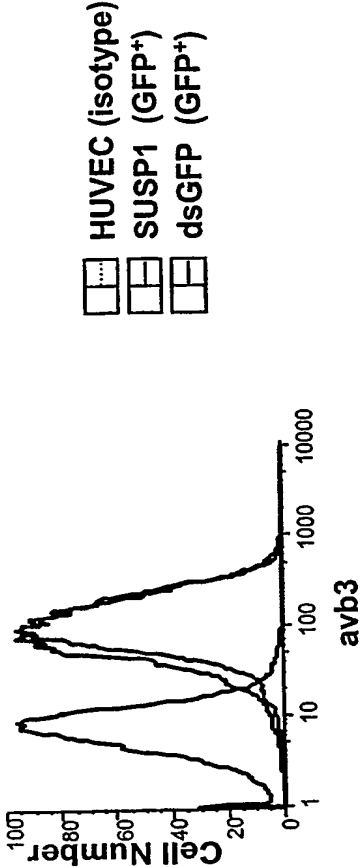


Figure 11



Sense orientation, N-terminal fusion

RECEIVED 23 FEB 2005

THIS PAGE BLANK DISPTOL

10/525575

GFP-SUSP-1 Screening Hit Does Not Affect Proliferation

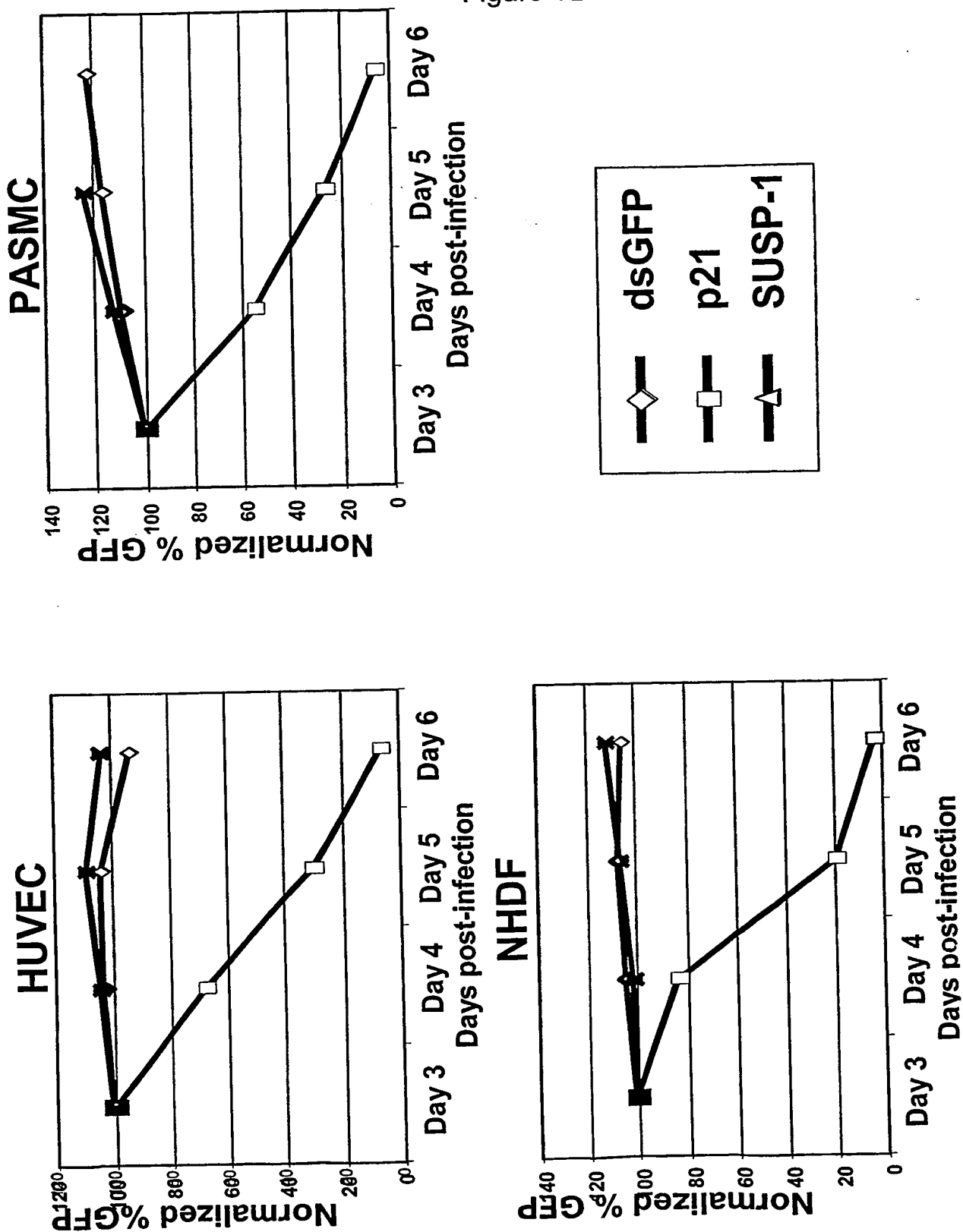


Figure 12

0115 Rec 01/PTO 2 3 FEB 2005

THIS PAGE BLANK (USPTO)

Expression Analysis of SUSP-1

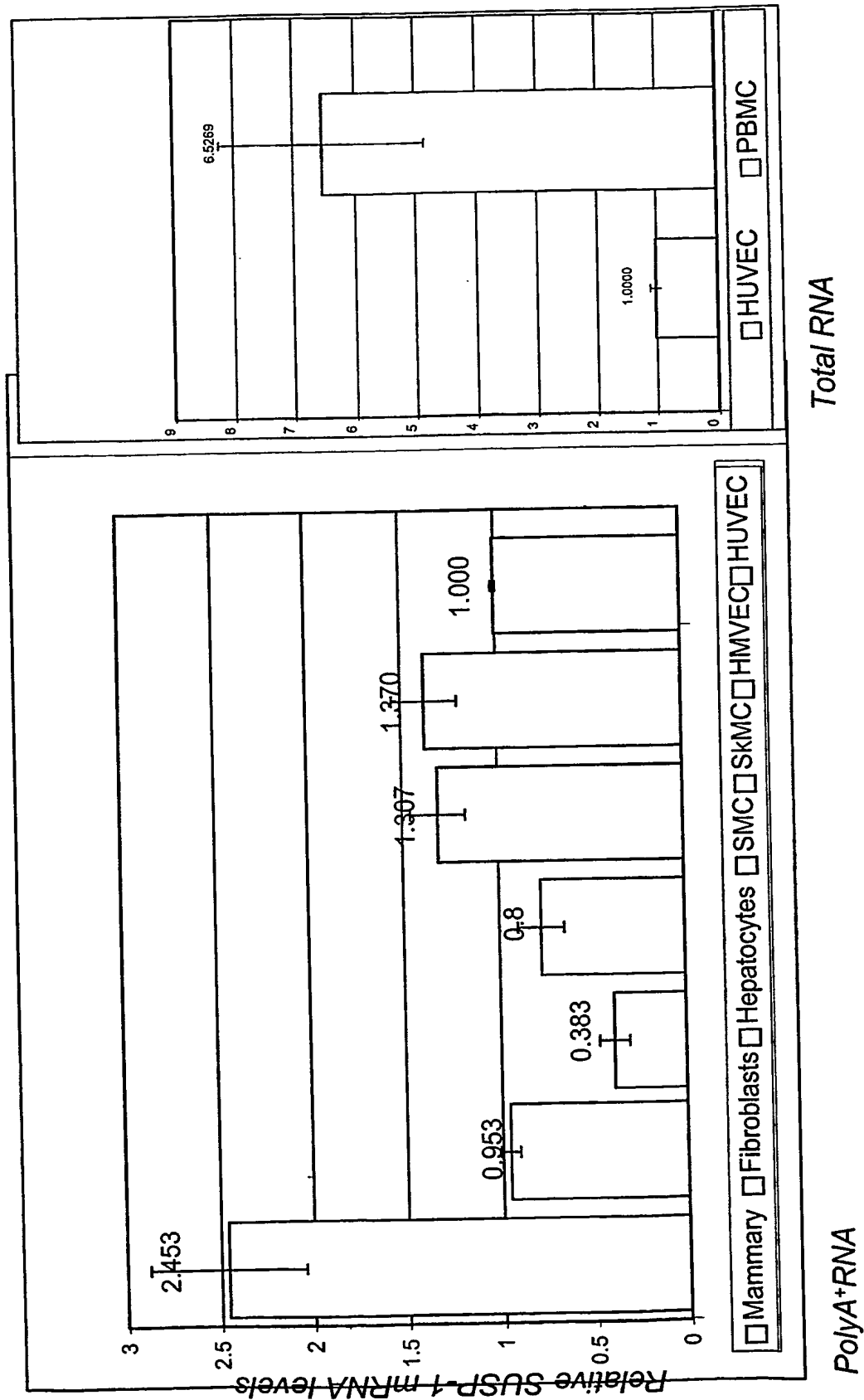


Figure 13

DT15 Rec/ST/PTO 23 FEB 2005

THIS PAGE BLANK (DISPTO)

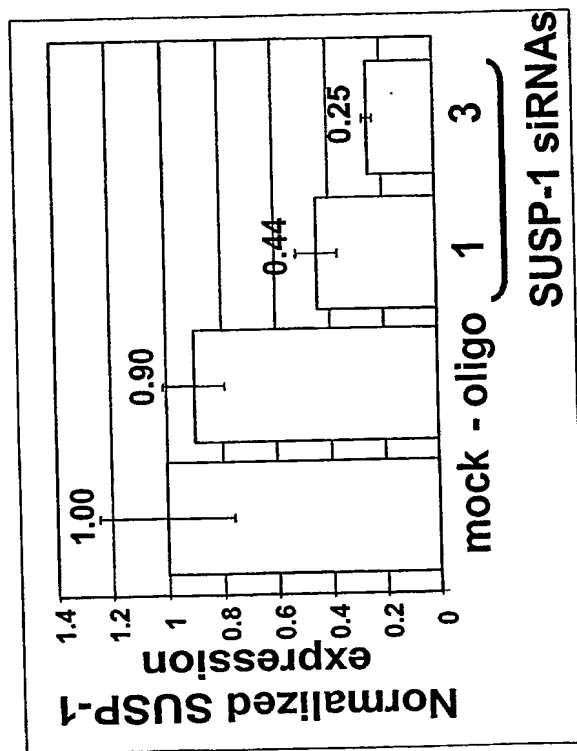
THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Figure 14

SUSP-1 siRNA inhibits SUSP-1 expression

Taqman



THIS PAGE BLANK (10/27/71)

Figure 15

A new sequence determination for the ABC transporter gene, SEQ ID NO:2.

ABC transporter>

GGACGCGCCTGGTGCCCCGGGGAGGGGCGCCACCGGGGAGGAGGAGGAG
GAGAAGGTGGAGAGGAAGAGACGCCCCCTCTGCCCCGAGACCTCTCAAGGC
CCTGACCTCAGGGGCCAGGGCACTGACAGGACAGGAGAGCAAGTTCTCTC
CACTTGGGCTGCCCCAAGAGGCCGCGACCCCTGGAGGGCCCTGAGCCCACC
GCACCAGGGGCCCCAGCACCCACCCCGGGGGCCTAAAGCGACAGTCTCAGG
GGCCATCGCAAGGTTTCCAGTTGCCTAGACAACAGGCCCAGGGTCAGAGC
AACAAATCCTTCCAGCCACCTGCCTCAACTGCTGCCCCAGGCACCAGCCCC
AGTCCCTACGCGGCAGCCAGCCAGGTGACATGCCGGTGCTCTCCAGGCC
CCGGCCCTGGCGGGGGAACACGCTGAAGCGCACGGCCGTGCTCCTGGCCC
TCGCGGCCTATGGAGCCACAAAGTCTACCCCTTGGTGCGCCAGTGCTTG
GCCCCGGCCAGGGGTCTTCAGGCGCCCGCGGGGAGCCACGCAGGAGGC
CTCCGGGGTTCGCGCGGCCAAAGCTGGCATGAACCGGGTATTCTTGCCAGC
GGCTCCTGTGGCTCCTGCGGCTGCTGTTCCCCCGGGTCTGTGCCGGGAG
ACGGGGCTGCTGGCCCTGCACTCGGCCCGCCTTGGTGAGCCGCACCTTCCT
GTGGGTGTATGTGGCCCGCCTGGACGGAAGGCTGGCCCGCTGCATCGTCC
GCAAGGACCCGCGGGCTTTTGGCTGGCAGCTGCTGCAGTGGCTCCTCATC
GCCCTCCCTGTACTCTCGTCAACAGTGCCATCCGTTACCTGGAGGGCCA
ACTGGCCCTGTCTGTTCCGCGAGCCGTCTGGTGGCCACGCCTACCGCCTCT
ACTTCTCCCAGCAGACCTACTACCGGGTCAGCAACATGGACGGGCGGCTT
CGCAACCCTGACCAGTCTCTGACGGAGGACGTGGTGGCCTTTGCGGCCTC
TGTGGCCACCTCTACTCAACCTGACCAAGCCACTCCTGGACGTGGCTG
TGACTTCTACACCCTGCTTCGGGCGGCCCGCTCCCGTGGAGCCGGCACA
GCCTGGCCCTCGGCCATCGCCGGCCTCGTGGTGTTCCTACGGCCAACGT
GCTGCGGGCCTTCTCGCCCAAGTTCGGGGAGCTGGTGGCAGAGGAGGCGC
GGCGGAAGGGGGAGCTGCGCTACATGCACTCGCGTGTGGTGGCCAACCTCG
GAGGAGATCGCCTTCTATGGGGGGCCATGAGGTGGAGCTGGCCCTGCTACA
GCGCTCCTACAGGACCTGGCCTCGCAGATCAACCTCATCCTTCTGGAAC
GCCTGTGGTATGTTATGCTGGAGCAGTTCCTCATGAAGTATGTGTGGAGC
GCCTCGGGCCTGCTCATGGTGGCTGTCCCATCATCACTGCCACTGGCTA
CTCAGAGTCAGATGCAAGAGCCGTGAAGAAGGCAGCCTTGGAAAAGAAGG
AGGAGGAGCTGGTGAGCGAGCGCACAGAAGCCTTCACTATTGCCCGCAAC
CTCCTGACAGCGGCTGCAGATGCCATTGAGCGGATCATGTCTGTCTACAA
GGAGGTGACGGAGCTGGCTGGCTACACAGCCCGGGTGACAGAGATGTTCC
AGGTATTGGAAGATGTTACGCGCTGTCACTTCAAGAGGCCAGGGAGCTA
GAGGACGCTCAGGCGGGGTCTGGGACCATAGGCCGTCTGGTGTCCGTGT
GGAGGGCCCCCTGAAGATCCGAGGCCAGGTGGTGGATGTGGAACAGGGGA
TCATCTGCGAGAACATCCCCATCGTCACGCCCTCAGGAGAGGTGGTGGT
GCCAGCCTCAACATCAGGGTGGAGGAAGGCATGCATCTGCTCATCACAGG
CCCCAATGGCTGCGGCAAGAGCTCCCTGTTCCGGATCCTGGGTGGGCTCT
GGCCACGTACGGTGGTGTGCTCTACAAGCCCCCAGCCAGCGCATGTTT
TACATCCCGCAGAGGCCCTACATGTCTGTGGGCTCCCTGCGTGACCAGGT
GATCTACCCGGACTCAGTGGAGGACATGCAAAGGAAGGGCTACTCGGAGC
AGGACCTGGAAGCCATCCTGGACGTCTGACCTGCACCACATCCTGCAG
CGGGAGGGAGGTTGGGAGGCTATGTGTGACTGGAAGGACGTCCTGTCTGGG
TGGCGAGAAGCAGAGAATCGGCATGGCCCGCATGTTCTACCACAGGCCCCA
AGTACGCCCTCCTGGATGAATGCACCAGCGCCGTGAGCATCGACGTGGAA
GGCAAGATCTTCCAGGCGGCCAAGGACGCGGGCATTGCCCTGCTCTCCAT
CACCCACCGGCCCTCCCTGTGGAAATACCACACACACTTGCTACAGTTTCG
ATGGGGAGGGCGGCTGGAAGTTCGAGAAGCTGGACTCAGCTGCCCGCCTG
AGCCTGACGGAGGAGAAGCAGCGGCTGGAGCAGCAGCTGGCGGGCATTCC
CAAGATGCAGCGGCGCCTCCAGGAGCTCTGCCAGATCCTGGGCGAGGCCG

THIS PAGE BLANK (USPTO)

50/525573

TGGCCCCAGCGCATGTGCCGGCACCTAGCCCGCAAGGCCCTGGTGGCCTC
CAGGGTGCCCTCCACCTGACACAACCGTCCCCGGCCCCCTGCCCCGCCCA
AGCTCGGATCACATGAAGGAGACAGCAGCACCCACCCATGCACGCACCCC
GCCCCTGCAATGCCCTGGCCCCCTCCTCCTAGAAAACCTTCCCGCCCTCGGG
AAAGTAGATGTGGAGGGTGGCGCCCTGCGTAACCTCGCCCTGTCCCTCC
CACTCCCTGGGGGCGCTGTTCACAGTGAAGTGGGCCCTGTCCAGGGCAGT
GAGTCCCTTACTTTGCTCCGTGGAGGAAGCTGGGGTACAAGGGGCCAGT
GCTGGCCACACAGCAGCGCAGCCGAGCCCCAGGAGCCCGTCAGGCCACAG
CCCCTGGCACTGCAGGTGGCCCTCCCTCCAGAGACTCGAGTCCCCATGATT
CCCTCCTCGTCAGTCTCTCAAAGACCCCATGGTCCATCCCCTGAGGGTGG
TCAGCCAAGGCTCCCGTTCCGTGGGATGCCATAAAAGCCGCCAGTGGGA
CCCACAGTCACACAGAGCGCCTCACCTGCATCCTCTCCCCACAAGAGCC
CCAAAGATCCCACGGGAGAGGGGAGAGGGACGCACAGCACTGCCTGCCAA
GCGAGAATGCAGGCCCCGCCCCCTCGGCCCTCACCACCTCTTTCTACAG
CCTAATTTATTGGATTCCCTATTTCGTAGCCATCTCCGTGGCCAATGTGAC
TACCGTGCCAGCAGCGGGGGCGGCCAGCCTCTGAGTCCCGTGGGGCCCC
GGCTCCCACCGGTGCCAAACCCAGCCCCCTGCGGCCGTACCCCGCCAGCC
TACACTGCCAGCCGCCACCGGGGCACACGGGCCTCTGCTTGCCAGCCAGG
AGTGCGGACACCATGTTCCAGCTCAGTGCCAAAGAGGGGTACCCAGGGG
GAGCTGTCTGCGGAGCCAGCGCTGCCCGAGAGAGACCCACCGCCACCG
TGTGCCTTTCCCGGGCCCTCAGCCCTCGGGCCGGGCACCAACCCAGTCC
CCCCAGTAAAAGCCTCCACTGGCAAAAAAAAAAAAAAAAAAAAAA

Figure 15 cont.

23 FEB 2005

THIS PAGE BLANK (USPTO)

Figure 16

A new sequence determination, SEQ ID NO:127 for HSPA5

SEQ ID NO:127

HSP5>

ATGGAGGAGGACAAGAAGGAGGACGTGGGCACGGTGGTTCGGCATCGACCT
GGGGACCACCTACTCCTGCGTCGGCGTGTTCAAGAACGGCCGCGTGGAGA
TCATCGCCAACGATCAGGGCAACCGCATCACGCCGTCTATGTCGCCTTC
ACTCCTGAAGGGGAACGTCTGATTGGCGATGCCGCAAGAACCAGCTCAC
CTCCAACCCCGAGAACACGGTCTTTGACGCCAAGCGGCTCATCGGCCGCA
CGTGGAATGACCCGTCTGTGCAGCAGGACATCAAGTTCTTGCCGTTCAAG
GTGGTTGAAAAGAAAATAAACCATAATTCAAGTTGATATTGGAGGTGG
GCAAACAAAGACATTTGCTCCTGAAGAAATTTCTGCCATGGTTCTCACTA
AAATGAAAGAAACCGCTGAGGCTTATTTGGGAAAGAAGGTTACCCATGCA
GTTGTTACTGTACCAGCCTATTTTAATGATGCCCAACGCCAAGCAACCAA
AGACGCTGGAATATTGCTGGCCTAAATGTTATGAGGATCATCAACGAGC
CTACGGCAGCTGCTATTGCTTATGGCCTGGATAAGAGGGAGGGGAGAAG
AACATCCTGGTGTGTTGACCTGGGTGGCGGAACCTTCGATGTGTCTCTTCT
CACCATTGACAATGGTGTCTTTCGAAGTTGTGGCCACTAATGGAGATACTC
ATCTGGGTGGAGAAGACTTTGACCAGCGTGTATGGAACACTTCATCAAA
CTGTACAAAAAGAAGACGGGCAAAGATGTCAGGAAAGACAATAGAGCTGT
GCAGAAACTCCGGCGCGAGGTAGAAAAGGCCAAACGGGGCCCTGTCTTCTC
AGCATCAAGCAAGAATTGAAATTGAGTCTTCTATGAAGGAGAAGACTTT
TCTGAGACCCTGACTCGGGCCAAATTTGAAGAGCTCAACATGGATCTGTT
CCGGTCTACTATGAAGCCCGTCCAGAAAGTGTGGAAAGATTCTGATTGA
AGAAGTCTGATATTGATGAAATTGTTCTTGTGGTGGCTCGACTCGAATT
CCAAAGATTCAAGCAACTGGTTAAAGAGTTCTTCAATGGCAAGGAACCATC
CCGTGGCATAAACCAGATGAAGCTGTAGCGTATGGTGTCTGTGTCCAGG
CTGGTGTGCTCTCTGGTGTATCAAGATACAGGTGACCTGGTACTGCTTGAT
GTATGTCCCCCTTACACTTGGTATTGAAACTGTGGGAGGTGTATGACCAA
ACTGATTTCCAAGGAACACAGTGGTGCCACCAAGAAGTCTCAGATCTTTT
CTACAGCTTCTGATAATCAACCAACTGTTACAATCAAGGTCTATGAAGGT
GAAAGACCCCTGACAAAAGACAATCATCTTCTGGGTACATTTGATCTGAC
TGGAATTCTCTCTGCTCCTCGTGGGGTCCACAGATTGAAGTCACCTTTG
AGATAGATGTGAATGGTATTCTTCGAGTGACAGCTGAAGACAAGGGTACA
GGGAACAAAAATAAGATCACAAATCACCATGACCAGAATCGCCTGACACC
TGAAGAAATCGAAAGGATGGTTAATGATGCTGAGAAGTTTGCTGAGGAAG
ACAAAAAGCTCAAGGAGCGCATTTGATACTAGAAATGAGTTGGAAAGCTAT
GCCTATTCTCTAAAGAATCAGATTGGAGATAAAGAAAAGCTGGGAGGTAA
ACTTTCTCTGAAGATAAGGAGACCATGGAAAAAGCTGTAGAAGAAAAGA
TTGAATGGCTGGAAAGCCACCAAGATGCTGACATTGAAGACTTCAAAGCT
AAGAAGAAGGAAGTGAAGAAATGTTCAACCAATTATCAGCAAACCTCTA
TGGAAGTGACAGGCCCTCCCCCAACTGGTGAAGAGGATACAGCAGAATCC
ACCACCACCACCACCAC

DTIC Ref ID: A61710 23 FEB 2005

THIS PAGE BLANK (USPTO)

Figure 17

Protein disulfide isomerase, SEQ ID NO:165.

SEQ ID NO:165

Novel protein disulfide isomerase

Novel (disulfide isomerase-like)

CGGACCAACACAGTATTGAGTCAACTGTGACCTTAAGATCAGAGGAACGTCAATACTGCCACAAGGCCACCTTTCCA
GAACTCGTGGGCAGGTAAACTATGCTTTGGATGTGCTTTCTTTCACCAAATCACTCAACTCAGGAGCCACAAATAG
TCCAGCAATTTCAATTTCCCTCAACGCTATTTTAGTCTCAAAGGAAACCATGTAAATTTTCATCAAGAGAAGGTCAAAG
GGGATATATCGCCACTGAAAATGTTTACACAGTGACCATGAGTTACACATTTACTTAGAGAACTTAACTTAATAAA
GAATCTGTAGAGTGTGTTGGCTTGGAAAACACACACACAAAGAAGATACCTCACGCTTAGTATGTTCTGCTTTCTGA
ACAGCCACCACTGGGAACCCAGTGGCCTCTGTGGGACTGAACTCCTAAACGCAGGGTGCGGGAGCTGGGCAGGAGAG
GTGACCTCCAACGTGTGTTCTAAAGTTTCGTCTTTTCGCTTGGCTCAGGACAAAGCGGTGTAACGAGTCAAAGTCTCTG
CCTCCACTGTGCTCACTGACTTTCTTCCCTCCTCGGAAAAGCAATAACGTGGGGTAGCCTCGT

DT15 Rec'd PST/PTO 23 FEB 2005

THIS PAGE BLANK (USPTO)

Figure 18

Chromosome 1 protein

SEQ ID NO:210

Clone 30 -Novel (Chromosome 1) >

ATCTACATGAGGTCCTGTAGATTGAGCAACTTTGGCTATTTGGGTAGCTGATCCACTGCCAAGCTGGCTCCCAGCTA
CCACCTGTTACTTTGACAACTCTAGCTTCAGGGCTACCAGGACCATTTGCAAACTTACACCCACAGAAAACTCAAA
GTCTCCAGAGGCATAAGCTCCAACCCCCAGAGAGCTCCACAGCATTCCTTCTCCTTCTAGTAGTACTTTTGGTTTCCA
GGAAAATCATCTCATGCTCGGGCCACAGTCACAGTCCTGACCAGCACCATAGAAAATTTAGCCATTATTCTAGTTGT
ATGGAGTCATTCAAAATACTAGTTCCGAAAGAAGTTTCAAAGGTCCAACGCCGAGCATCTTTTACCTTTGCTAATTT
TTCCATAAATTGAAAGCCTT

THIS PAGE RI ANK 118701

Figure 19

Chromosome 3 protein

SEQ ID NO:218

Clone 95 -Novel (Chromosome 3/ H41) >

GCGTCGCTCGGCGTTAGCCAAGGCCCCGGGCGGCCACCCTCCGGGGGCACTAGGTCTGGGGCCGCAGTGCCCAGCA
CAGAGCAGCGTTTATCGGGACGAAGATGAATGGAAAGAATTGGAGCAAAAAGAGGTTGATTACAGCGGCCTCAGGGG
TTCAGGCAATGCAAATAAGCAGTGAAAAGGAAGAAGACGATAATGAAAAG

THIS PAGE BLANK (USPTO)

Figure 20

Chromosome 17

SEQ ID NO:227

Clone 147 - Novel (Chromosome 17) >

CTGCTTCTCTTCTAGGATAGTTTCCCTCTAGAAATCCATGTTGCTCCTTCTCTCATTAAATGATCAGATATTTATGT
GCCTCCTTAGCTAAGAGGCCTTCCTTATCTACCCCTAACTAAGAGAATCCCCATATCACTCTTATCCTTTTATCCTGA
TTTATTTTTCCATAAACTTGCCACATTACTGGGCATTATCTATCATTTAGTAATGTGTATATTGCTTATTGACAGT
CTTCCCTACTATAATATGAGCTCCATGAGGCAGAGACTTTGTTTTGTTCACTGCTGAATTTCCAGCACCTACAACAA
GCCTACTCACTAAACATTTGTTGAATTAATCTTCAGGATACATCAAATGTAAAATATCCTATCATATATTTATTGAT
GTGATGCATAGGAAAATGTCTGGAAGGGTACATACCAAAATGTTAATTATTTATTTAACCTGGTGGGAGGTGGTTTT
TCTTTCCCTTTTCTTTTCTTTTGGAGACAGAGTTTCGCTCTCGTTGCCAGGCTGGAGTGGAATGGCAAG

RECEIVED

23 FEB 2005

THIS PAGE BLANK (USPTO)

Figure 21

Chromosome 8 protein

SEQ ID NO:232

Novel (chromosome 8) >

GCTTGCAAAGGAGAGGCTGTGACTACCAAGTCGTGTCAACAACCTGAATGGCTGAAATACCCAACTTGCCCATGCAA
ATGGGCTTGGGTCTCTCCTGGCAGCCGCTTTGAAGGCTCTAGACTTATCTGTGAACCTCTTTTGTGAGAGGGTCTT
TCCAACCTAGTGGTTTATTCTTTGACTCTCCTCATACTTTTTTGCCAGAGAGTGAGAGTGAGAAGGGAGGGCTAATG
CCTGAGCTCCTGCCCTTTCTATGCAGTGAGGGTCAAGATCCTCAGCTAGTGTGAGGGAACTGGTGTAACCTGGGT
CTCTCATTTTCTACCATCCAAGTTGCC

DT15 Rec'D OCT/PTO 23 FEB 2005

THIS PAGE BLANK (USPTO)

Figure 22

Chromosome 9 protein

SEQ ID NO:248

Clone 49 -Novel (Chromosome 9) >

CTGCAGGAGACCACAGGCAGGTGCCCACCTAAGAGGGACAGCCACAGAAACCCTCTAACCTCAGCACTGCACTCCAC
CACGACCACCCACGCAGGCCCTCAGCACCGCACTCCACCACGAACACCCACACAGGCTGAGGCTGGAGCCAGAAG
CTGCCAGAACATGGGACCACAGGGCCAGGCACCACACAAACATGGCGACGGACACAGCCATCCAACCCGACTCGGAC
CTCCGCCAGGCCCCAGCGCACAACCATCTGGGATCCCCAGGAAAAGAGCTGCGTGCGGCCAG

THIS PAGE BLANK (USPTO)

Figure 23

1/226 protein

SEQ ID NO:274

Clone 1/226 - Novel (Clone1/226) >

GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTCGCTCTTC
TTGCTTGTCGTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTCTCCGCCC
CAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACTTCACTTCATCCCTTCTACCCA
TCACAACATACAAAGCAACTACACCTGGATTTTTTCAAACAACTTTTATTTCTCAGAGTCTTCCTTAATCCTATGG
AACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATATAAAAATATAG
ACTTTTAAAAAAAAAAAAAAAAAAAAA

THIS PAGE BLANK (USPTO)

Figure 24

A new sequence determination for FLJ10688

SEQ ID NO:285

Novel (FLJ10688) >

CACCCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCCAGAAGCTGCCAGCACATCTCTGCCTCAGAAG
CCACACTTGAAGTTAGCACGCGTTCAGAGTCAAAATGGCATAGTACTGTCATGGAGTGCCTGGAGGTGGATCGAAG
CTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACAATGGAAAA
AGATTGGGGAAGTCAAGGCACTTCCCTTGCCCATGGCATGTACTCTACCCAGTTTGTATCTGGTAGCAAATACTAC
TTTG

THIS PAGE BLANK (USPTO)

Figure 25

A new sequence determination for KIAA1583

SEQ ID NO:307

Clone 89 -Novel (KIAA1583)>

ACTGCAGGTGGCAGCCACGGGCGCGCCGCTCGGCCTCATCTGACGCCTCTGCAGCGGGTTCCGCAGGCCTGCAGGGC
GGGGAGGCCGGGACTGGCCGTGACGCTGAACGGCCCAGCCTGCCAGGGCCCAGCTGCTGGAGACCCGCAGCTCGT
CCCCGGCGGGTCCCTAATCACCAGCAGCTCCTGTTTCTCAAAACGCAGACATCCGCCCCCTCTTGGGGTCAGGCCCTTC
CACCTTGCAGGCGAGCCGCCCCAGCCCACTCCCGACTGGCGCTATGCCTCGATCACCGCTCTTGCTCCCAAGTGGAC
CGCAGGGGAGACGCTCTCTTACGGGGACCCTGGGGGCGCTCACTCTCTGAAGGGCCTGGAAGCTAGATTCCAGAGGC
GTGGGCCACCTCTCCCTGGGTTTTGGGGAGCCCCCTCCGAGGGTGTTTCATTTCTGAGCTCTGTGTCATCTTAGGCT
CTGAGGGT

23 FEB 2005

THIS PAGE BLANK (USPTO)

Figure 26

A new sequence determination for KIAA1814

SEQ ID NO:317

Clone 75 -Novel (KIAA1814) >

CTGGACGGCCTGGCTGGGCTGAAGGGCGAGGACAGCCGAGCAAGGAGGCAGGGGAGGGCGGCCTACCGCTGTGCGG
GCCCACGGACAAGACCCCACTGCTGAGCGGCAAGGCCGCCAAGGCCCCGGGACCGCGAGGTCGACCTCAAGAATGGCC
ACAACCTCTTCATCTCTGCGGCGGCCGTGCCTCCCGGAAGCCTCCTCAGCGGCCCGGCCTGGCCCCGGCGGCCTCCT
CCGCAGGCGGCGCGGCGTCCCTCCGCCCAGACGCACCGGTCTTCTGGGCCCCCTTCCCGCCGGGACCGCAGTTCGCG
CTCGGCCCCATGTCCCTGCAGGCCAACCTCGGCTCCGTGGCCGGCTCCTCCGTGCTGCAGTCGCTGTTTCAGCTCTGT
GCCGGCCGCGCAGGCCTGGTGCACGTGTCGTCCGCTGCCACCAGACTGACCAACTCGCACGCCATGGGCAG

THIS PAGE BLANK (USPTO)

Figure 27

Chromosome 4 protein

SEQ ID NO:320

Novel (maps to chromosome 4)

Novel (Chromosome 4)>

GGAATAAAAGAGTGGAAATGGGGATTTCCAGGTGCTCCCCTGGTTCATCTAGGCACCAGAGAGCTGCACTAGCAGGT
CTATCATGAATCTCCTTGGAATGCTCATTTTTAGTCCTACTTGATGTGTCTGTTTCTGGAAATGCAGTATTTTAAAT
GTATCTCAACAAAATATTTTATGATTAGTAAGCTTATTCTTATATAAAGGACAATTTTTTTCCTTTTTCACAGGTTC
TAATAATTTTTTATTTAATAATTAGATCTATTAGATTTTATTCATAACTGTGGTAGTTGAAGTACCTTCTAAGCTGA
GTTTCAATTTGAGAATAAACCTTGGGGTATCATTACAGAAAATTTTGTCTCAATCTGCTTTGTATTTGAAAGATATG
AGATTCTTGAATTATATATCTTACAGACTAGTCCCCAAAAGAATACGTGTTTCCTTACCTTTAATTTCTCATGGTAG
TTAGTCTGTGAAT

DT15 Res PCT/PTO 23 FEB 2005

TUIC DAGE RI ANK (ISPTM)

Figure 28

A new sequence determination for peroxidasin/melanoma antigen related protein

SEQ ID NO:323

Novel peroxidasin-like/ melanoma antigen>

CGACCTGGCCAGCCACCGCGGCCTGCTGCGGCAGGGCATCGTGCAGCGGTCCGGGAAGCCGCTGCTCCCCCTTCGCCA
CCGGGCGCGCCACGGAGTGCATGCGGGACGAGAACGAGAGCCCCATCCCCTGCTTCCTGGCCGGGACCACCGCGCC
AACGAGCAGCTGGGCCTGACCAGCATGCACACGCTGTGGTTCCGCGAGCACAACCGCATTGCCACGGAGCTGCTCAA
GCTGAACCCGCACTGGGACGGCGACACCATCTACTATGAGACCAGGAAGATCGTGGGTGCGGAGATCCAGCACATCA
CCTACCAGCACTGGCTCCCGAAGATCCTGGGGGAGGTGGGCATGAGGACGCTGGGAGAGTACCACGGCTACGACCCC
GGCATCAATGCTGGCATCTTCAACGCCTTCGCCACCGCGGCCTTCAGGTTTGGCCACACGCTTGTCAACCCACTGCT
TTACCGGCTGGACGAGAACTTCCAGCCCATTGCACAAGATCACCTCCCCCTTCACAAAGCTTTCTTCTCTCCCTTCC
GGATTGTGAATGAGGGCGGCATCGATCCGCTTCTCAGGGGGCTGTTCCGGGGTGGCGGGGAAAATGCGTGTGCCCTCG
CAGCTGCTGAACACGGAGCTCACGGAGCGGCTGTTCTCCATGGCACACACGGTGGCTCTGGACCTGGCGGCCATCAA
CATCCAGCGGGGCCGGGACCACGGGATCCCACCCTACCACGACTACAGGGTCTACTGCAATCTATCGGCGGCACACA
CGTTCGAGGACCTGAAAAATGAGATTAAAAACCTTGAGATCCGG

THIS PAGE BLANK (USPTO)

Figure 29

WD40/SOCS box protein

SEQ ID NO:329

Novel WD40/SOCS box protein>

CGGCTCTGTTGCCGTGCAAGTGGAATCATCTTACCTCTAAGACATGACTTTCACATGCCAGGGGAGAGAAAGATCCA
TGACTAGTACACTGGCAATCTGGTTTTGCTACATTCTATTCACAATCCCAAAGAAATGCTATTTCAATGCAAGACCA
GATGTTTGGCCCATTTATCCAGCAACTCCCTTTGACAGGACGATTTACCTTGCTACCAAAGAAGCACAAGATGTGGT
GTTGCTTAAAAAAGTCCTGTATGTGAGGAACTCTTTCATTTTCTTGGGGATTGGCAGTGCTAGGACTTGGTAAGTTG
TTAGGAACTTCGAAGGGCTTCCGGCATAAGTGCTTCAGTGAGGACAGGACCCTAGGAGCTGTCCAGAACTGGACG
TGGCCATCTCTTGTCCCTGTGGCAATGACTCCACCATGTGGAAAAAATGTGCAGCAAAGCCCATTGGTCATAGGAGC
AAATGCAATGGGAGTTTTTCAGTTCCAGGGCCCAGATCCTGAGGAGTCTGTCATCTGCCACCGTGGCAAGGTACAAGC
CCCCG

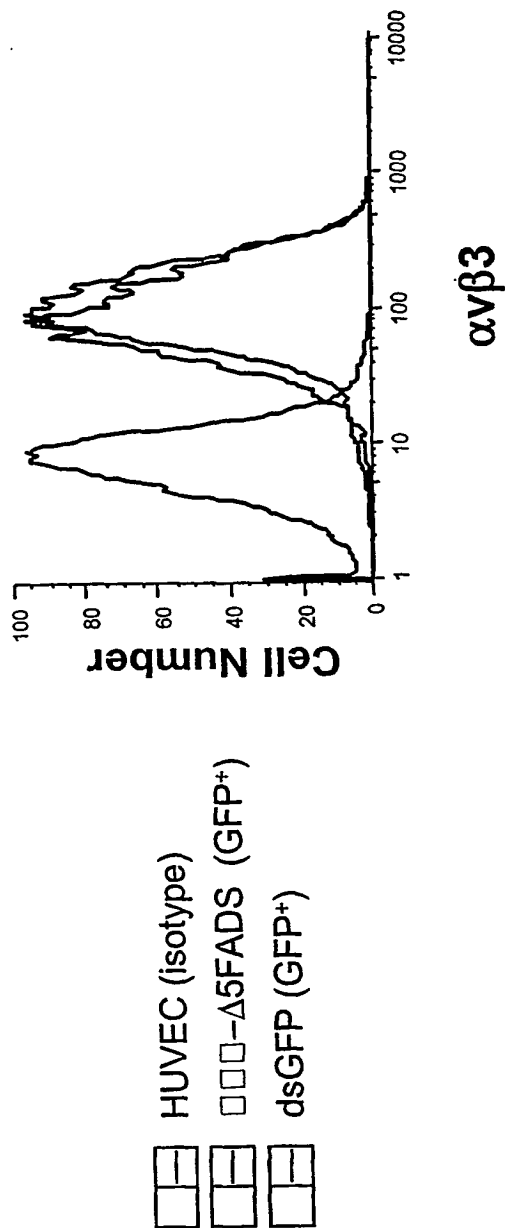
23 FEB 2005

THIS PAGE BLANK (USPTO)

GFP-Δ5FADS Antisense Screening Hit Downregulates Surface αvβ3 Levels



• GFP- Δ5FADS screening hit: 300 bp antisense



DATE 23 FEB 2005

THIS PAGE RI ANK 1150PM

GFP- $\Delta 5$ FADS Antisense Screening Hit does not Affect Cell Proliferation

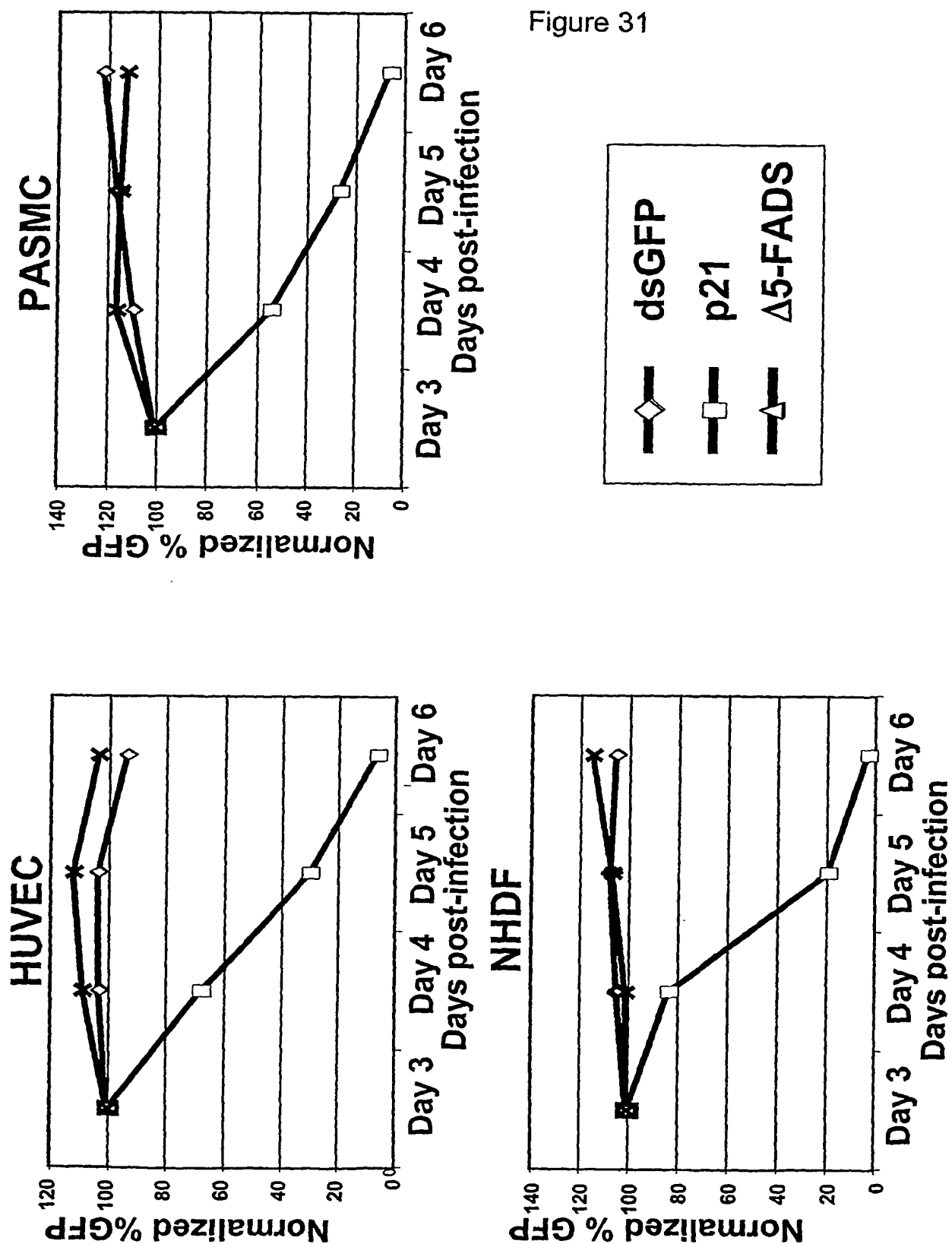


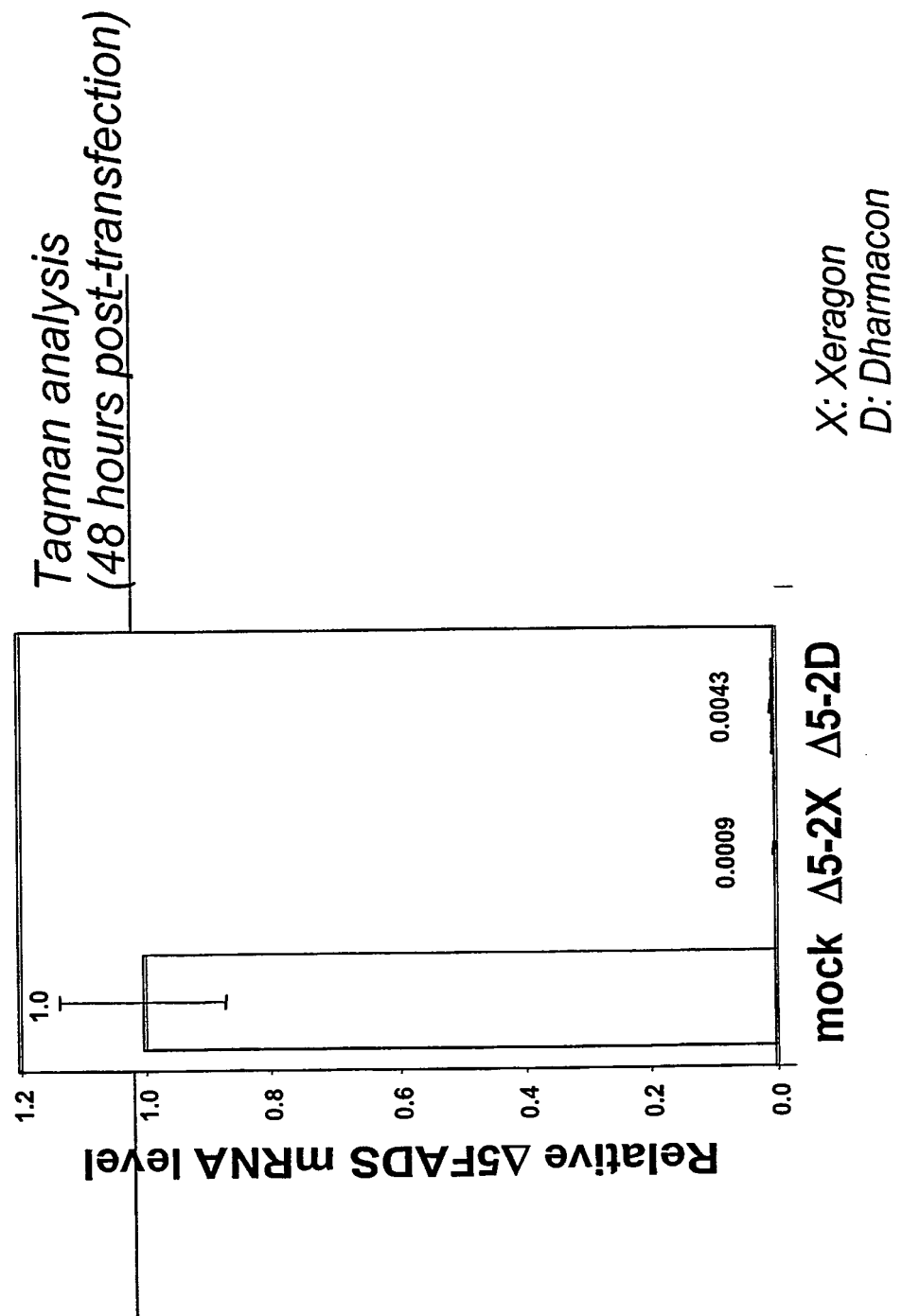
Figure 31

THIS PAGE BLANK (USPTO)

10/525573

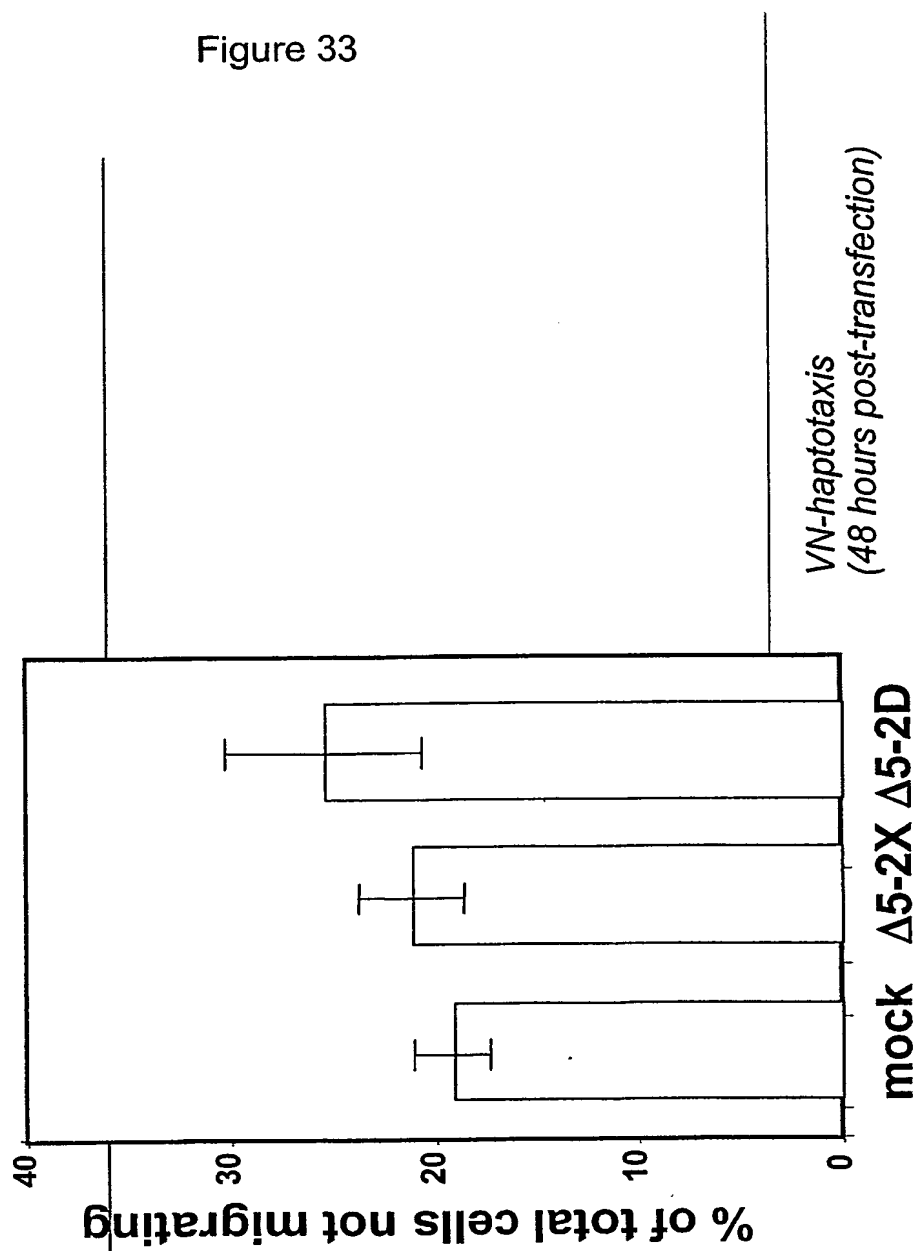
Figure 32

siRNA($\Delta 5-2$) Blocks $\Delta 5$ FADS Expression in HUVEC



THIS PAGE IS BLANK

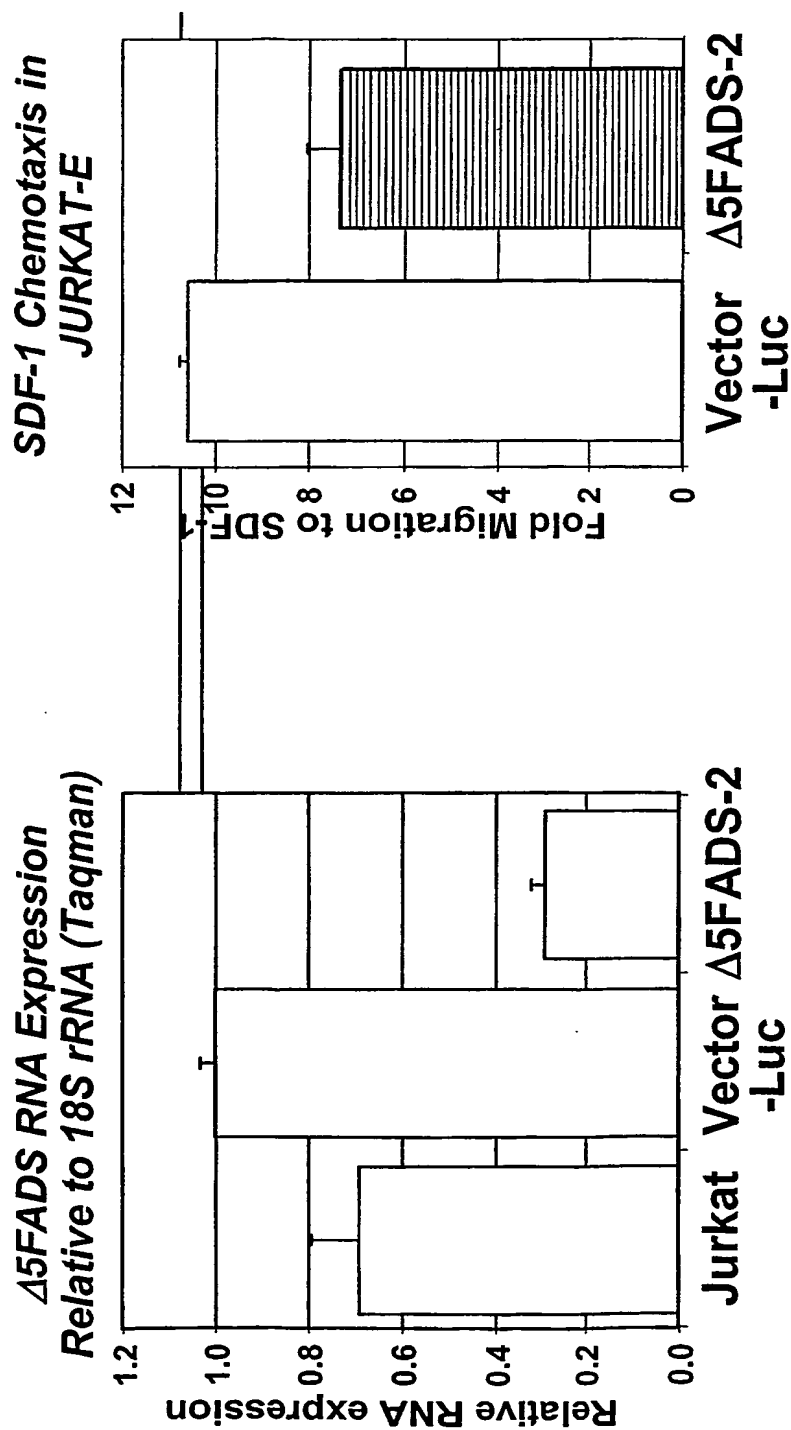
$\Delta 5$ FADS RNAi Inhibits Haptotaxis



THIS PAGE BLANK (USPTO)

Figure 34

Δ5-FADS RNAi Vector Reduces SDF-1 Induced Chemotaxis in Jurkat Cells



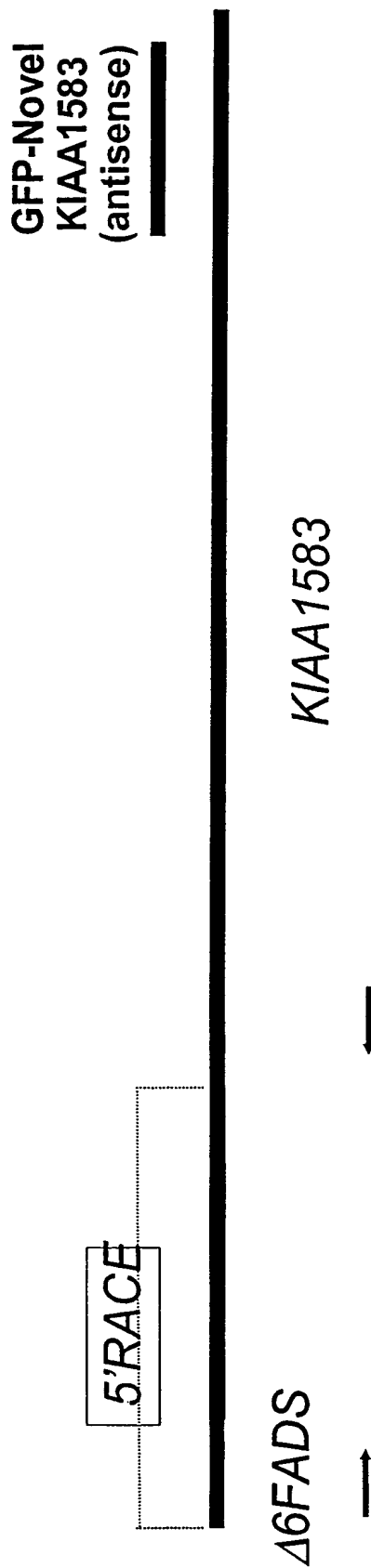
THIS PAGE BLANK (USPTO)

23 FEB 2003

Figure 35

Δ 6FADS is a Functional Screening Hit

- 5' RACE analysis indicates that the Δ 6FADS transcript is spliced with the Novel (KIAA1583) transcript.
- Δ 6FADS and KIAA1583 sequences are both located on Chromosome 11q12



*Splicing independently
validated by RT-PCR
and sequence analysis*

034

23 FEB 2005

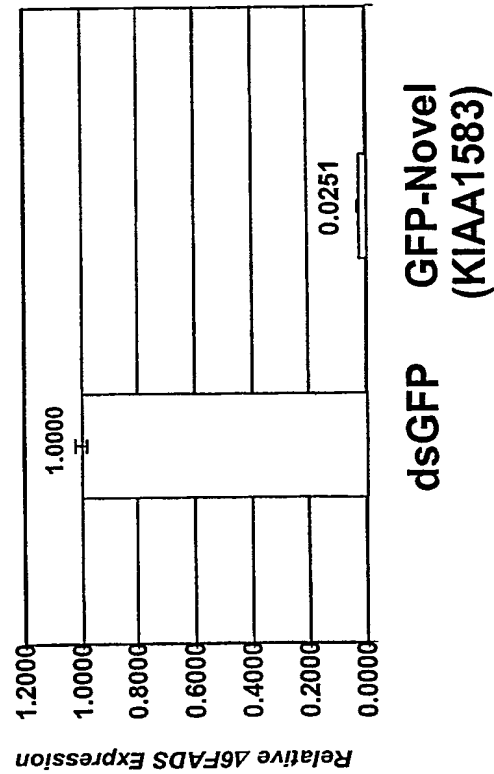
THIS PAGE IS ANK

Figure 36

$\Delta 6$ FADS is a Functional Screening Hit

- HUVECs expressing the GFP- Novel (KIAA1583) screening hit have strongly reduced levels of $\Delta 6$ FADS mRNA.

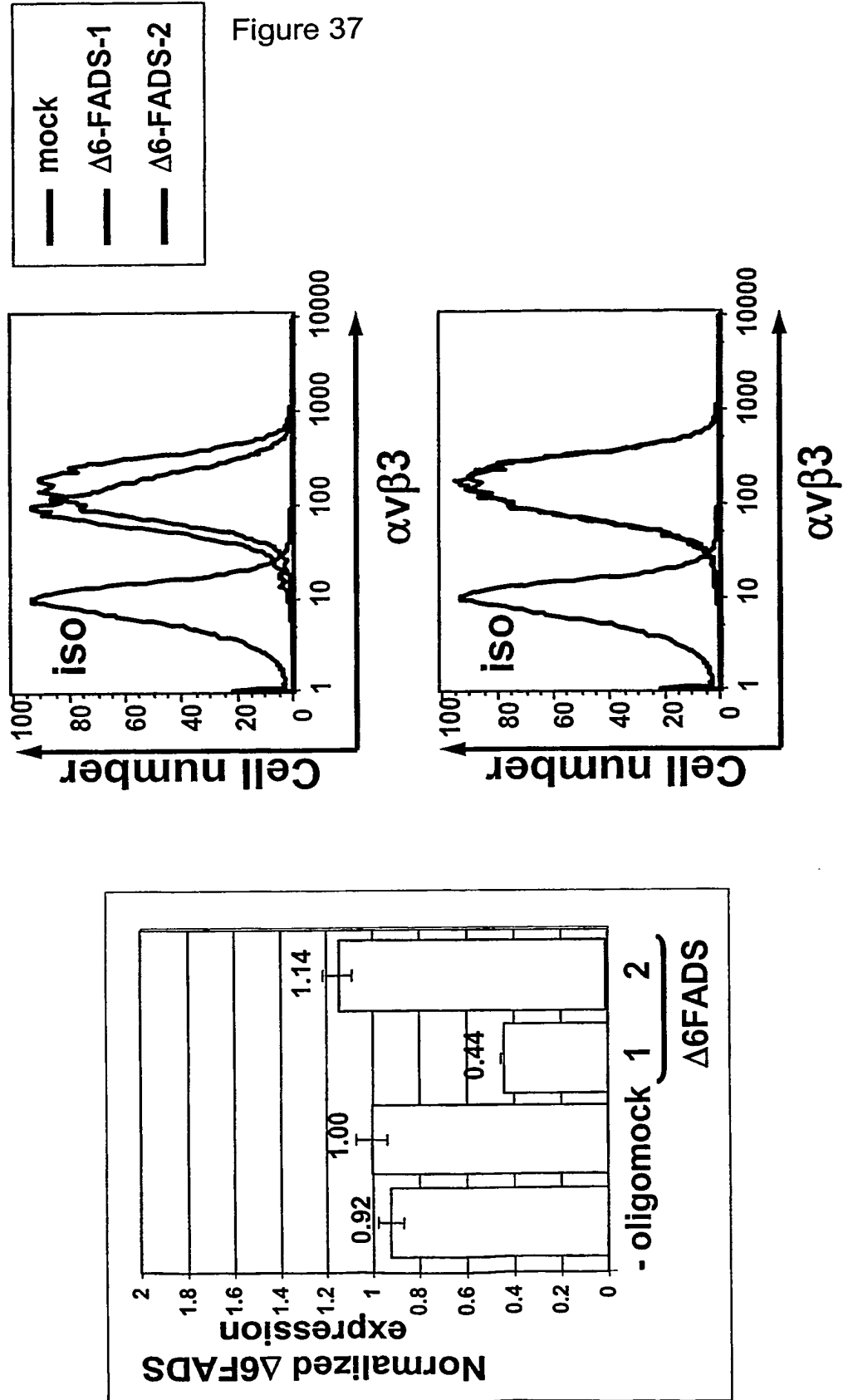
Taqman analysis of $\Delta 6$ FADS mRNA levels



DTIC ***** PUT/PTC 23 FEB 2005

THIS PAGE RI ANK *****

$\Delta 6$ FADS RNAi Reduces $\alpha v\beta 3$ Surface Expression



THIS PAGE BLANK (11/07/05)

BRIEF SUMMARY OF THE INVENTION

The present application identifies, for the first time, a number of proteins and DNA molecules involved in regulation of angiogenesis, *e.g.*, angiogenesis regulatory proteins and DNA molecules. The invention further relates to methods for identifying and using agents, including small organic molecules, antibodies, peptides, cyclic peptides, nucleic acids, antisense nucleic acids, RNAi, and ribozymes, that modulate angiogenesis via modulation of angiogenesis regulatory proteins and DNA molecules; as well as to the use of expression profiles and compositions in diagnosis and therapy of diseases related to insufficient or increased angiogenesis.

In one aspect, the present invention provides a method for identifying a compound that modulates angiogenesis, the method comprising the steps of:

(i) contacting the compound with an angiogenesis regulatory nucleic acid, or an angiogenesis regulatory polypeptide or a fragment thereof encoded by a nucleic acid, wherein the nucleic acid hybridizes under stringent conditions to a second nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523; and (ii) determining the functional effect of the compound upon the nucleic acid or polypeptide.

In one embodiment, the functional effect is determined in vitro. In another embodiment, the functional effect is a physical effect. In another embodiment, the functional effect is determined by measuring ligand or substrate binding to the polypeptide. In another embodiment, the functional effect is a chemical effect. In another embodiment, the functional effect is determined by measuring an enzymatic activity.

THIS PAGE BLANK (USPTO)

In one embodiment, the polypeptide is expressed in a eukaryotic host cell. In a further embodiment, the functional effect is a physical effect. In another embodiment, the functional effect is determined by measuring ligand or substrate binding to the polypeptide. In a further embodiment, the functional effect is a chemical or phenotypic effect. In another embodiment, the functional effect is determined by measuring an enzymatic activity. In another embodiment, the host cell is an endothelial cell. In a further embodiment, the functional effect is determined by measuring $\alpha v \beta 3$ expression or haptotaxis, or chemotaxis, or co-culture tube formation.

In one embodiment, modulation is inhibition of angiogenesis.

In one embodiment, the polypeptide is recombinant. In another embodiment, the nucleic acid comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523.

In another embodiment, the polypeptide comprises a sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID

THIS PAGE BLANK (USPTO)

NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:520, SEQ ID NO:522, and SEQ ID NO:524.

In one embodiment, the compound is an antibody, an antisense molecule, a small organic molecule, or a peptide.

5 In another aspect, the present invention provides a method for identifying a compound that modulates angiogenesis, the method comprising the steps of (i) contacting the compound with a nucleic acid, or a polypeptide or a fragment thereof encoded by a nucleic acid, wherein the nucleic acid hybridizes under stringent conditions to a second nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:3, SEQ
10 ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID
15 NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID
20 NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523; (ii) determining the functional effect of the compound upon the nucleic acid or polypeptide; and (iii) expressing the nucleic acid or polypeptide in a cell, contacting the nucleic acid or polypeptide with the compound, and determining the
25 phenotypic or chemical effect upon the cell.

In another aspect, the present invention provides a method of modulating angiogenesis in a subject, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified as a modulator of angiogenesis using the methods described herein. In one embodiment, the subject is a human. In a further
30 embodiment, the compound is an antibody, an antisense molecule, a small organic molecule, a peptide, or an RNAi molecule. In another embodiment, the compound inhibits angiogenesis.

In another aspect, the present invention provides a method of modulating angiogenesis in a subject, the method comprising the step of administering to the subject a

THIS PAGE BLANK (USPTO)

therapeutically effective amount of a polypeptide, the polypeptide encoded by a nucleic acid that hybridizes under stringent conditions to a second nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:63, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:165, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:218, SEQ ID NO:232, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:340, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, and SEQ ID NO:498.

In another aspect, the present invention provides a method of modulating angiogenesis in a subject, the method comprising the step of administering to the subject a therapeutically effective amount of a nucleic acid that hybridizes under stringent conditions to a second nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523.

In another aspect, the present invention provides the nucleic acid sequences of SEQ ID NO:165, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, and SEQ ID NO:330.

THIS PAGE BLANK (USPTO)

In another aspect, the present invention provides the amino acid sequences of SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, and SEQ ID NO:331.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides a schematic diagram of an assay for modulators of angiogenesis.

Figure 2 provides results of an experiment demonstrating the effect of a plakoglobin-GFP fusion protein expression on levels of the cell surface marker $\alpha\beta3$.

Plakoglobin is an exemplar angiogenesis regulatory protein. Human umbilical vein endothelial (HUVEC) cells were transfected with a vector expressing the plakoglobin-GFP fusion protein ((GGS)₂ = SEQ ID NO:525) or a control vector, expressing GFP only. Cells were incubated with APC-labeled antibodies directed against the cell surface marker $\alpha\beta3$. The X-axis depicts cell number and the Y-axis depicts the amount of $\alpha\beta3$ -APC antibody derived fluorescence. Cells transfected with the plakoglobin-GFP fusion protein construct exhibit lower $\alpha\beta3$ expression levels than control cells.

Figure 3 also provides results of an experiment demonstrating the effect of a plakoglobin-GFP fusion protein (*i.e.*, GFP- Δ N-plakoglobin) expression on levels of the cell surface marker $\alpha\beta3$. The figure also depicts the presence of armadillo repeats in the functional screen hit, GFP- Δ N-plakoglobin.

Figure 4 demonstrates that GFP- Δ N-Plako expression downregulates vitronectin receptors ($\alpha\beta3$ and $\alpha\beta5$) but not the fibronectin receptor ($\alpha5\beta1$).

Figure 5 demonstrates that expression of GFP- Δ N-Plakoglobin downregulates endogenous catenin levels, as well as expression of endogenous plakoglobin and VE-cadherin.

Figure 6 demonstrates that GFP- Δ N-Plakoglobin expression upregulates surface levels of the laminin receptor, $\alpha6\beta4$.

Figure 7 provides results of an experiment demonstrating the effect of a HoxB2-GFP fusion protein expression (*i.e.*, Δ N-HoxB2) on levels of the cell surface marker $\alpha\beta3$. Cells transfected with the Δ N-HoxB2-GFP fusion protein construct exhibit lower $\alpha\beta3$ expression levels than control cells. The figure also provides the structure of the functional screen hit, demonstrating that the homeodomain is present in Δ N-HoxB2.

Figure 8 demonstrates that Δ N-HoxB2 downregulates both α v and $\beta3$ integrin subunit surface expression, as well as $\alpha\beta5$ surface expression.

THIS PAGE BLANK (USPTO)

Figure 9 demonstrates that Δ N-HoxB2 Inhibits Proliferation in HUVECs, a primary cell line.

Figure 10 demonstrates that full length HoxB2 phenocopies Δ N-HoxB2 in the downregulation of α v β 3.

Figure 11 provides results of an experiment demonstrating the effect of a SUSP-1 GFP fusion protein expression (*i.e.*, GFP-SUSP-1) on levels of the cell surface marker α v β 3. Cells transfected with the GFP-SUSP-1 fusion protein construct exhibit lower α v β 3 expression levels than control cells. The figure also provides the structure of the functional screen hit, demonstrating that GFP-SUSP-1 is an N-terminal fusion in the sense orientation and that a cDNA insert is present before the wild-type start codon.

Figure 12 demonstrates that the GFP-SUSP-1 screening hit does not affect proliferation when expressed in HUVEC (human umbilical vein endothelial cells), PASMC (pulmonary artery smooth muscle cells), and NHDF (normal human dermal fibroblasts).

Figure 13 provides expression analysis of SUSP-1 mRNA in a variety of cell lines.

Figure 14 demonstrates that SUSP-1 siRNAs inhibit SUSP-1 mRNA expression.

Figure 15 provides a new sequence determination for a nucleic acid encoding an ABC transporter protein (SEQ ID NO:2).

Figure 16 provides a new sequence determination for a nucleic acid encoding an HSPA5 protein (SEQ ID NO:127).

Figure 17 provides a new sequence determination for a nucleic acid encoding a protein disulfide isomerase (SEQ ID NO:165).

Figure 18 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 1 (SEQ ID NO:210).

Figure 19 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 3 (SEQ ID NO:218).

Figure 20 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 17 (SEQ ID NO:227).

Figure 21 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 8 (SEQ ID NO:232).

Figure 22 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 9 (SEQ ID NO:248).

THIS PAGE BLANK (USPTO)

Figure 23 provides a new sequence determination for a nucleic acid encoding a 1/226 protein (SEQ ID NO:274).

Figure 24 provides a new sequence determination for a nucleic acid encoding an FLJ10688 protein (SEQ ID NO:285).

5 Figure 25 provides a new sequence determination for a nucleic acid encoding a KIAA1583 protein (SEQ ID NO:307).

Figure 26 provides a new sequence determination for a nucleic acid encoding a KIAA1814 protein (SEQ ID NO:317).

10 Figure 27 provides a new sequence determination for a nucleic acid encoding a novel protein from human chromosome 4 (SEQ ID NO:320).

Figure 28 provides a new sequence determination for a nucleic acid encoding a peroxidasin/melanoma antigen related protein (SEQ ID NO:323).

Figure 29 provides a new sequence determination for a nucleic acid encoding a WD40/SOCS box protein (SEQ ID NO:329).

15 Figure 30 provides provides results of an experiment demonstrating the effect of a $\Delta 5FADS$ antisense screening hit (*i.e.*, GFP- $\Delta 5FADS$ or D5ADS) on levels of the cell surface marker $\alpha\beta 3$. Cells transfected with the GFP- $\Delta 5FADS$ antisense construct exhibit lower $\alpha\beta 3$ expression levels than control cells. The figure also provides the structure of the functional screen hit, demonstrating that GFP- $\Delta 5FADS$ is in the antisense orientation and
20 that a the anti sense portion of the $\Delta 5FADS$ nucleic acid is about 300 base pairs long.

Figure 31 demonstrates that the GFP-D5FADS antisense screening hit does not affect cell proliferation.

Figure 32 demonstrates that $\Delta 5FADS$ siRNAs inhibit $\Delta 5FADS$ mRNA expression.

25 Figure 33 demonstrates that $\Delta 5FADS$ siRNAs inhibit haptotaxis.

Figure 34 demonstrates that $\Delta 5FADS$ siRNAs inhibit SDF-1 induced chemotaxis in Jurkat cells.

Figure 35 demonstrates that $\Delta 6FADS$ is a functional screening hit. The figure shows that $\Delta 6FADS$ transcript is spliced to the KIAA1583 functional screen hit.

30 Figure 36 demonstrates tha $\Delta 6FADS$ mRNA levels are straongly reduced in cells that express the antisense KIAA1583 functional screen hit.

Figure 37 that $\Delta 6FADS$ siRNAs inhibit $\Delta 6FADS$ mRNA expression and also reduce $\alpha\beta 3$ surface expression.

THIS PAGE BLANK (USPTO)

DETAILED DESCRIPTION OF THE INVENTION

Introduction

Using a functional assay, angiogenesis regulatory proteins and nucleic acids have been identified and cloned from a green fluorescent protein (GFP) fusion library.

5 Primary endothelial cells were transduced with retroviral cDNA/GFP fusion libraries. Transduced cells were selected and assayed for alterations consistent with inhibition of angiogenesis, e.g. downregulation of cell surface expression of $\alpha v \beta 3$ or downregulation of haptotaxis.

10 Using the screen described above, fifty-two proteins and/or DNA molecules were shown for the first time to regulate angiogenesis. Thirty-four of the proteins and/or DNA molecules are known, while fifteen of the proteins and/or DNA molecules are completely novel. However, this is the first demonstration of a function in regulation of angiogenesis for these molecules.

15 The ABC transporter is involved in peroxisome biogenesis. Relevant sequence data for the protein, nucleic acids encoding the ABC transporter, and related sequences include the nucleic acid accession number NM_000033, SEQ ID NO:3; protein accession number NP_000024, SEQ ID NO:4; a new sequence determination, SEQ ID NO:2; and the functional screen hit, SEQ ID NO:1. The protein has an ATP-binding domain. The gene encoding the ABC transporter is linked to a genetic disease, adrenoleukodystrophy
20 (Mosser *et al.*, *Nature* 361:682-3 (1993)). An antisense version of the ABC transporter gene, e.g., SEQ ID NO:1, was identified as a functional hit in the screen for modulators of angiogenesis.

25 PCB1, a poly (rC) binding protein functions as a translational co-activator and also binds 3' UTRs of RNA, most likely through KH domains, thereby regulating RNA stability. Relevant sequence data for the protein, nucleic acids encoding PCB1, and related sequences include the nucleic acid accession number NM_006196, SEQ ID NO:32; protein accession number NP_006187, SEQ ID NO:33; and the functional screen hit, SEQ ID NO:5-31. An antisense version of the gene encoding PCB1 was identified as a functional hit in the screen for modulators of angiogenesis.

30 SLC1A5, an amino acid transporter, has ten transmembrane spanning segments. Relevant sequence data for the protein, nucleic acids encoding SLC1A5, and related sequences include the nucleic acid accession number NM_005628, SEQ ID NO:43; protein accession number NP_005619, SEQ ID NO:44; and the functional screen hit, SEQ ID

THIS PAGE BLANK (USPTO)

NO:34-42. An antisense version of the gene encoding SLC1A5 was identified as a functional hit in the screen for modulators of angiogenesis.

Chromobox homolog 6 is involved in chromatin regulation. Relevant sequence data for the protein, nucleic acids encoding chromobox homolog 6, and related sequences include the nucleic acid accession number NM_014292, SEQ ID NO:57; protein accession number NP_055107, SEQ ID NO:58; and the functional screen hit, SEQ ID NO:45-56. An antisense version of chromobox homolog 6 was identified as a functional hit in the screen for modulators of angiogenesis.

Cytochrome C oxidase subunit 1 is encoded by a mitochondrial gene and transfers electrons from cytochrome C to O₂ (Ingman *et al.*, *Nature* 408:708-713 (2000); (Leninger, *Principles of Biochemistry* (1984); Stryer, *Biochemistry* (1995)) Relevant sequence data for the protein, nucleic acids encoding cytochrome C oxidase, and related sequences include the nucleic acid accession number NC_001807, SEQ ID NO:63; protein accession number NP_536845, SEQ ID NO:64; and the functional screen hit, SEQ ID NO:59-62. A sense version of the gene encoding cytochrome C oxidase, encoding a 21 amino acid peptide, was identified as a functional hit in the screen for modulators of angiogenesis.

Δ -5 fatty acid desaturase (FADSD5 or Δ 5FADS) introduces double bonds into a fatty acyl chain. The protein is involved in arachadonic acid synthesis and has a fatty acid desaturase domain and a heme-binding domain. Relevant sequence data for the protein, nucleic acids encoding Δ -5 fatty acid desaturase, and related sequences include the nucleic acid accession number NM_013402, SEQ ID NO:68; protein accession number NP_037534, SEQ ID NO:69; and the functional screen hit, SEQ ID NO:65-67. An antisense version of Δ -5 fatty acid desaturase gene was identified as a functional hit in the screen for modulators of angiogenesis.

The dynactin p27 subunit protein was shown to be an angiogenesis regulatory protein. The protein has an RGD domain. Relevant sequence data for the protein, and the nucleic acid encoding dynactin p27 subunit protein, include the nucleic acid accession number NC_006571, SEQ ID NO:70; and protein accession number NP_006562.1, SEQ ID NO:71. An antisense version of the gene encoding dynactin p27 subunit protein was identified as a functional hit in the screen for modulators of angiogenesis.

Elongation factor 1 alpha (EF1 α) was shown to be involved in regulation of angiogenesis. Relevant sequence data for the protein, nucleic acids encoding EF1 α , and related sequences include the nucleic acid accession number NC_001402, SEQ ID NO:76;

THIS PAGE BLANK (USPTO)

functional screen hit. Relevant sequence data for the protein, nucleic acids encoding HSPA5, and related sequences include the nucleic acid accession number NM_005347, SEQ ID NO:128; protein accession number NP_005338, SEQ ID NO:129; a new sequence determination, SEQ ID NO:127; and the functional screen hit, SEQ ID NO:122-126.

5 Interferon gamma receptor 1 (IFN γ R1) is a cytokine receptor that induces antiangiogenic genes. Relevant sequence data for the protein, nucleic acids encoding IFN γ R1, and related sequences include the nucleic acid accession number NM_000416, SEQ ID NO:139; protein accession number NP_000407, SEQ ID NO:140; and the functional screen hit, SEQ ID NO:130-138. An antisense version of the IFN γ R1 gene was identified as
10 a functional hit in the screen for modulators of angiogenesis.

Importin α 4 (karyopherin α 3) is a subunit of the nuclear localization signal receptor. Relevant sequence data for the protein, nucleic acids encoding importin α 4, and related sequences include the nucleic acid accession number NM_002267, SEQ ID NO:153; protein accession number NP_002258.1, SEQ ID NO:154; and the functional screen hit, SEQ
15 ID NO:141-152. An antisense version of the importin α 4 gene was identified as a functional hit in the screen for modulators of angiogenesis.

The lysosomal pepstatin-insensitive protease (CLN2) is encoded by a gene associated with juvenile neuronal ceroid lipofuscinosis disease. Relevant sequence data for the protein, nucleic acids encoding CLN2, and related sequences include the nucleic acid
20 accession number NM_000391, SEQ ID NO:163; protein accession number NP_000382, SEQ ID NO:164; and the functional screen hit, SEQ ID NO:155-162. An antisense version of the CLN2 gene was identified as a functional hit in the screen for modulators of angiogenesis.

A novel protein disulfide isomerase was shown to have a role in angiogenesis.
25 Relevant sequence data for the nucleic acid encoding the novel protein disulfide isomerase, includes SEQ ID NO:165. An antisense version of gene encoding the novel protein disulfide isomerase was identified as a functional hit in the screen for modulators of angiogenesis. The antisense hit encodes a 45 or 32 amino acid peptide that may also be useful as a therapeutic or to modulate angiogenesis.

30 Microtubule-associated protein 4 (MAP4) binds tubulin and may link microtubules to other elements of the cytoskeleton. See *e.g.*, Chapin & Bulinski, *J. Cell. Sci* 98:27-36 (1991); and Chapin *et al.*, *Biochemistry* 34:2289-2301 (1995). Four isoforms of MAP4 are known. Relevant sequence data for the proteins, nucleic acids encoding MAP4 isoforms, and related sequences include the nucleic acid accession numbers NM_002375,

THIS PAGE BLANK (USPTO)

Plasminogen activator inhibitor 1 (SERPINE1 or PAI-1) was identified as an angiogenesis regulatory protein. PAI-1 is a protease inhibitor. Relevant sequence data for the protein, nucleic acids encoding PAI-1, and related sequences include the nucleic acid accession number NM_000602, SEQ ID NO:419; protein accession number NP_000593, SEQ ID NO:420; and the functional screen hit, SEQ ID NO:410-418. The functional screen hit is in the antisense orientation.

Proteosomal subunit Y, an INF γ regulated proteosomal subunit, was identified as an angiogenesis regulatory protein. Relevant sequence data for proteosomal subunit Y includes the functional screen hit, SEQ ID NO:421. The functional screen hit is in the antisense orientation.

Rap2B was identified as an angiogenesis regulatory protein. Rap2B is a member of the RAS family of GTPases. Relevant sequence data for the protein, nucleic acids encoding Rap2B, and related sequences include the nucleic acid accession number XM_003032, SEQ ID NO:428; protein accession number XP_003032, SEQ ID NO:429; and the functional screen hit, SEQ ID NO:422-427. The functional screen hit is in the antisense orientation.

Semaphorin 3F was identified as an angiogenesis regulatory protein. Semaphorin 3F is a ligand for NP2. Relevant sequence data for the protein, nucleic acids encoding semaphorin 3F, and related sequences include the nucleic acid accession number U38276, SEQ ID NO:437; protein accession number AAB18276, SEQ ID NO:438; and the functional screen hit, SEQ ID NO:430-436. The functional screen hit is in the antisense orientation.

SPARC was identified as an angiogenesis regulatory protein. Relevant sequence data for the protein, nucleic acids encoding SPARC, and related sequences include the nucleic acid accession number NM_003118, SEQ ID NO:439; and protein accession number NP_003109, SEQ ID NO:440.

ssDNA binding protein-3 was identified as an angiogenesis regulatory protein. Relevant sequence data for the protein, nucleic acids encoding ssDNA binding protein-1, and related sequences include the nucleic acid accession number NM_018070, SEQ ID NO:445; protein accession number NP_060540, SEQ ID NO:446; and the functional screen hit, SEQ ID NO:441-444. The functional screen hit is in the sense orientation.

Sumo protease (SUSP-1) was identified as an angiogenesis regulatory protein. Relevant sequence data for the protein, nucleic acids encoding SUSP-1, and related sequences include the nucleic acid accession number NM_015571, SEQ ID NO:456; protein

THIS PAGE BLANK (USPTO)

The terms “angiogenesis regulatory protein or nucleic acid” or a nucleic acid encoding “angiogenesis regulatory protein” refer to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino acid sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acids, to a polypeptide encoded by a nucleic acid of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, and SEQ ID NO:519, or an amino acid sequence of SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:520, SEQ ID NO:522, and SEQ ID NO:524 or other sequences listed herein; (2) specifically bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence of SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ

THIS PAGE BLANK (USPTO)

ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:520, SEQ ID NO:522, and SEQ ID NO:524, immunogenic fragments thereof, and conservatively modified variants thereof; (3)

specifically hybridize under stringent hybridization conditions to a nucleic acid encoding
SEQ ID NO:4, SEQ ID NO:33, SEQ ID NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID NO:520, SEQ ID NO:522, and SEQ ID NO:524, e.g., a nucleic acid sequence of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523; or its complement, and conservatively modified variants thereof; (4) have a nucleic acid sequence that has greater than about 95%, preferably greater than about 96%, 97%, 98%, 99%, or higher nucleotide

THIS PAGE RI ANK (USPTO)

sequence identity, preferably over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ ID NO:523, or the complement of any of those sequences. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or any mammal. The nucleic acids and proteins of the invention include both naturally occurring or recombinant molecules. When available, accession numbers for the human angiogenesis regulatory proteins and genes are provided.

The phrase “functional effects” in the context of assays for testing compounds that modulate activity of an angiogenesis regulatory proteins includes the determination of a parameter that is indirectly or directly under the influence of an angiogenesis regulatory protein polypeptide, e.g., a chemical or phenotypic effect such as loss-of angiogenesis phenotype represented by a change in expression of a cell surface marker, such as $\alpha v\beta 3$ integrin, or changes in cellular proliferation, especially endothelial cell proliferation; or enzymatic activity, or, e.g., a physical effect such as ligand binding or inhibition of ligand binding. A functional effect therefore includes ligand binding activity, the ability of cells to proliferate, expression in cells undergoing angiogenesis, and other characteristics of angiogenic cells. “Functional effects” include *in vitro*, *in vivo*, and *ex vivo* activities.

Angiogenesis assays are described, e.g., in *Angiogenesis Protocols* (Murray, ed., 2001).

By “determining the functional effect” is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of an angiogenesis regulatory proteins, e.g., measuring physical and chemical or phenotypic effects. Such functional effects can be measured by any means known to those skilled in the

THIS PAGE BLANK (USPTO)

family, the integrin family, the selectin family, and the like; *see, e.g.,* Pigott & Power, *The Adhesion Molecule Facts Book I* (1993). Similarly, toxins and venoms, viral epitopes, hormones (e.g., opiates, steroids, etc.), intracellular receptors (e.g. which mediate the effects of various small ligands, including steroids, thyroid hormone, retinoids and vitamin D; peptides), drugs, lectins, sugars, nucleic acids (both linear and cyclic polymer configurations), oligosaccharides, proteins, phospholipids and antibodies can all interact with various cell receptors.

Synthetic polymers, such as polyurethanes, polyesters, polycarbonates, polyureas, polyamides, polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides, and polyacetates can also form an appropriate tag or tag binder. Many other tag/tag binder pairs are also useful in assay systems described herein, as would be apparent to one of skill upon review of this disclosure.

Common linkers such as peptides, polyethers, and the like can also serve as tags, and include polypeptide sequences, such as poly Gly sequences of between about 5 and 200 amino acids (SEQ ID NO:526). Such flexible linkers are known to persons of skill in the art. For example, poly(ethylene glycol) linkers are available from Shearwater Polymers, Inc. Huntsville, Alabama. These linkers optionally have amide linkages, sulfhydryl linkages, or heterofunctional linkages.

Tag binders are fixed to solid substrates using any of a variety of methods currently available. Solid substrates are commonly derivatized or functionalized by exposing all or a portion of the substrate to a chemical reagent which fixes a chemical group to the surface which is reactive with a portion of the tag binder. For example, groups which are suitable for attachment to a longer chain portion would include amines, hydroxyl, thiol, and carboxyl groups. Aminoalkylsilanes and hydroxyalkylsilanes can be used to functionalize a variety of surfaces, such as glass surfaces. The construction of such solid phase biopolymer arrays is well described in the literature. *See, e.g.,* Merrifield, *J. Am. Chem. Soc.* 85:2149-2154 (1963) (describing solid phase synthesis of, e.g., peptides); Geysen *et al.*, *J. Immun. Meth.* 102:259-274 (1987) (describing synthesis of solid phase components on pins); Frank & Doring, *Tetrahedron* 44:60316040 (1988) (describing synthesis of various peptide sequences on cellulose disks); Fodor *et al.*, *Science*, 251:767-777 (1991); Sheldon *et al.*, *Clinical Chemistry* 39(4):718-719 (1993); and Kozal *et al.*, *Nature Medicine* 2(7):753759 (1996) (all describing arrays of biopolymers fixed to solid substrates). Non-chemical approaches for fixing tag binders to substrates include other common methods, such as heat, cross-linking by UV radiation, and the like.

THIS PAGE BLANK (USPTO)

AKIVLRRHLSQDHTVPCFAASELHRSRAEHTKENGLPYQSPSPVPGSMKLSPOBPRPLSPGALQLAGEKSS
 EKGLRERAYGSSGELITSLPISIPPLSTVQPNKLPVSIPLASVVLPSRAERARSTPSPVLQPRDPSSTLEK
 QIGANAHGAGSRSLALAPAGFSYAGSVAISGALAGSPASLTPGAEPATLDESSSSGSLFATVGSRSSTPO
 HPLLLAQPRNSLPASPAHQLSSSPRLGGAAQGPLPEASKGDLPSDSGFSDPESEAKRRIVFTITTGAGSA
 5 KQSPSSKHSPLTASARGDCVPSHGDSTRRRGRKRASAGTPSLSAGVSPKRRALPSVAGLFTQPSGSPLN
 LNSMVSININQPLEITAISSPETSLSKSSPVYQDHDQPPVLKKERPLSQTNGAHYSPLTSDEEPGSEDEPS
 SARIERKIATISLESKSPPKTLENGGGLAGRKPAPAGEPVNSSKWKSTFSPISDIGLAKSADSPLQASSA
 LSQNSLFTFRPALEEPSADAKLAAHPRKGFPGSLSGADGLSPGTNPANGCTFGGGLAADLSLHSFSDGAS
 10 LPHKGPEAAGLSSPLSFPSQRGKEGSDANPFLSKRQLDGLAGLKGEGRGKEAGEGGLPLCGPTDKTPLL
 SGKAAKARDREVDLKNHNLFIASAAVPPGSLLSGPGPLAPAASSAGGAASSAQTHRSFLGPFPPGPQFAL
 GPMSLQANLGSVAGSSVLSLFSVPAAGLVHVSSAATRLTNSHAMGSFSGVAGGTGGVFNHVPSPAS
 AHPFGARVGRGAACGSATLGPSPLQAAASASASSFQAPASVETRPPPPPPPPPLPPPAHLGRSPAGPP
 VLHAPPPPNAAALPPPPTLLASNPEPALLQSLASLPPNQAFLPPTSAASLPPANASLSIKLTSLPHKGARP
 15 SFTVHHQPLPRLALAAQAPGIPQASATGPSAVVWSLGMPPPYAAHLSGVKPR

SEQ ID NO:320

Novel (maps to chromosome 4)

Novel (Chromosome 4) >

GGAATAAAAGAGTGGAATGGGGATTTCCAGGTGCTCCCCTGGTTCATCTAGGCACCAGAGAGCTGCACTAGCAG
 20 GTCTATCATGAATCTCCTTGGAATGCTCATTTTATGCTCTACTTGATGTGTCTGTTTCTGGAAATGCAGTATTTT
 TAATGTATCTCAACAAAATATTTTATGATTAGTAAGCTTATTCTTATATAAAGGACAATTTTTTCTCTTTTTCAC
 AGGTTCTAATAATTTTTTATTTAATAATTAGATCTATTAGATTTTATTCATAACTGTGGTAGTTGAAGTACCTTC
 TAAGCTGAGTTTCAGATTTGAGAATAAACCTTGGGGTATCATACAGAAAATTTTGTCTCAATCTGCTTTGTATTT
 25 GAAAGATATGAGATCTTGAATTATATCTTACAGACTAGTCCCCAAAAGAATACGTGTTTCTTACCTTTAAT
 TTCTCATGGTAGTTAGTCTGTGAAT

SEQ ID NO:321

Novel peroxidase-like/ melanoma antigen

>GL2-86-4M13F Direction: sense

CGACCTGGCCAGCCACCGCGGCTGCTGCGGCAGGGCATCGTGACGCGGTCCGGGAAGCCGCTGCTCCCCCTTCGC
 30 CACCGGGCCGCCACGGAGTGCATGCGGGACGAGAACGAGAGCCCCATCCCCTGCTTCTGGCCGGGGACCACCG
 CGCCAACGAGCAGCTGGGCCTGACCAGCATGCACACGCTGTGGTTCCGCGAGCACAACCGCATTGCCACGGAGCT
 GCTCAAGCTGAACCCGCACTGGGACGGCGACACCATCTACTATGAGACCAGGAAGATCGTGGGTGCGGAGATCCA
 GCACATCACCTACCAGCACTGGCTCCCGAAGATCCTGGGGGAGGTGGGCATGAGGACGCTGGGAGAGTACCACGG
 35 CTACGACCCCGGCATCAATGCTGGCATCTTCAACGCCTTCGCCACCGCGGCCTTCAGGTTTGGCCACACGCTTGT
 CAACCCACTGCTTTACCGGCTGGACGAGAACTTCCAGCCCATTGACACAAGATCACCTCCCCCTTCACAAAGCTTT
 CTTCTCTCCCTTCCGGATTG

SEQ ID NO:322

>GL2_93_2B08_G3F1 Direction: sense

CGACCTGGCCAGCCACCGCGGCTGCTGCGGCAGGGCATCGTGACGCGGTCCGGGAAGCCGCTGCTCCCCCTTCGC
 40 CACCGGGCCGCCACGGAGTGCATGCGGGACGAGAACGAGAGCCCCATCCCCTGCTTCTGGCCGGGGACCACCG
 CGCCAACGAGCAGCTGGGCCTGACCAGCATGCACACGCTGTGGTTCCGCGAGCACAACCGCATTGCCACGGAGCT
 GCTCAAGCTGAACCCGCACTGGGACGGCGACACCATCTACTATGAGACCAGGAAGATCGTGGGTGCGGAGATCCA
 45 GCACATCACCTACCAGCACTGGCTCCCGAAGATCCTGGGGGAGGTGGGCATGAGGACGCTGGGAGAGTACCACCG
 CTACGACCCCGGCATCAATGCTGGCATCTTCAACGCCTTCGCCACCGC

SEQ ID NO:323

Novel peroxidase-like/ melanoma antigen >

CGACCTGGCCAGCCACCGCGGCTGCTGCGGCAGGGCATCGTGACGCGGTCCGGGAAGCCGCTGCTCCCCCTTCGC
 50 CACCGGGCCGCCACGGAGTGCATGCGGGACGAGAACGAGAGCCCCATCCCCTGCTTCTGGCCGGGGACCACCG
 CGCCAACGAGCAGCTGGGCCTGACCAGCATGCACACGCTGTGGTTCCGCGAGCACAACCGCATTGCCACGGAGCT
 GCTCAAGCTGAACCCGCACTGGGACGGCGACACCATCTACTATGAGACCAGGAAGATCGTGGGTGCGGAGATCCA
 GCACATCACCTACCAGCACTGGCTCCCGAAGATCCTGGGGGAGGTGGGCATGAGGACGCTGGGAGAGTACCACGG
 55 CTACGACCCCGGCATCAATGCTGGCATCTTCAACGCCTTCGCCACCGCGGCCTTCAGGTTTGGCCACACGCTTGT
 CAACCCACTGCTTTACCGGCTGGACGAGAACTTCCAGCCCATTGACACAAGATCACCTCCCCCTTCACAAAGCTTT
 CTTCTCTCCCTTCCGGATTGTGAATGAGGGCGGCATCGATCCGCTTCTCAGGGGGCTGTTCCGGGTGGCGGGGAA
 AATGCGTGTGCCCCCGCAGCTGCTGAACACGAGCTCACGGAGCGGCTGTTCTCATGGCACACACGGTGGCTCT
 60 GGACCTGGCGGCCATCAACATCCAGCGGGGCGGGACACGGGATCCACCCCTACCACGACTACAGGGTCTACTG
 CAATCTATCGGCGGCACACACGTTTCGAGGACCTGAAAAATGAGATTAAAAACCTGAGATCCG

THIS PAGE BLANK (USPTO)

WHAT IS CLAIMED IS:

1 1. A method for identifying a compound that modulates angiogenesis,
2 the method comprising the steps of:

3 (i) contacting the compound with a nucleic acid, or a polypeptide or a
4 fragment thereof encoded by a nucleic acid, wherein the nucleic acid hybridizes under
5 stringent conditions to a second nucleic acid comprising a nucleotide sequence selected
6 from the group consisting of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID
7 NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID
8 NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID
9 NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID
10 NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID
11 NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID
12 NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID
13 NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID
14 NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID
15 NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID
16 NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID
17 NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID
18 NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ
19 ID NO:523 ; and

20 (ii) determining the functional effect of the compound upon the nucleic
21 acid or polypeptide.

1 2. The method of claim 1, wherein the functional effect is determined
2 *in vitro*.

1 3. The method of claim 2, wherein the functional effect is a physical
2 effect.

1 4. The method of claim 2, wherein the functional effect is determined
2 by measuring ligand or substrate binding to the polypeptide.

1 5. The method of claim 2, wherein the functional effect is a chemical
2 effect.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

1 ID NO:76, SEQ ID NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID
2 NO:128, SEQ ID NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID
3 NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID
4 NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID
5 NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID
6 NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID
7 NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID
8 NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID
9 NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID
10 NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID
11 NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID
12 NO:521, and SEQ ID NO:523.

1 19. The method of claim 1, wherein the polypeptide comprises a
2 sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:33, SEQ ID
3 NO:44, SEQ ID NO:58, SEQ ID NO:64, SEQ ID NO:69, SEQ ID NO:71, SEQ ID
4 NO:77, SEQ ID NO:82, SEQ ID NO:87, SEQ ID NO:90, SEQ ID NO:121, SEQ ID
5 NO:129, SEQ ID NO:140, SEQ ID NO:154, SEQ ID NO:164, SEQ ID NO:170, SEQ ID
6 NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:184, SEQ ID NO:203, SEQ ID
7 NO:287, SEQ ID NO:298, SEQ ID NO:309, SEQ ID NO:319, SEQ ID NO:325, SEQ ID
8 NO:331, SEQ ID NO:341, SEQ ID NO:352, SEQ ID NO:366, SEQ ID NO:378, SEQ ID
9 NO:385, SEQ ID NO:407, SEQ ID NO:409, SEQ ID NO:420, SEQ ID NO:429, SEQ ID
10 NO:438, SEQ ID NO:440, SEQ ID NO:446, SEQ ID NO:457, SEQ ID NO:463, SEQ ID
11 NO:482, SEQ ID NO:485, SEQ ID NO:494, SEQ ID NO:497, SEQ ID NO:499, SEQ ID
12 NO:520, SEQ ID NO:522, and SEQ ID NO:524.

1 20. The method of claim 1, wherein the compound is an antibody.

1 21. The method of claim 1, wherein the compound is an antisense
2 molecule.

1 22. The method of claim 1, wherein the compound is a small organic
2 molecule.

1 23. The method of claim 1, wherein the compound is a peptide.

THIS PAGE BLANK (USPTO)

1 24. A method for identifying a compound that modulates angiogenesis,
2 the method comprising the steps of:

3 (i) contacting the compound with a nucleic acid, or a polypeptide or a
4 fragment thereof encoded by a nucleic acid, wherein the nucleic acid hybridizes under
5 stringent conditions to a second nucleic acid comprising a nucleotide sequence selected
6 from the group consisting of SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:43, SEQ ID
7 NO:57, SEQ ID NO:63, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:76, SEQ ID
8 NO:81, SEQ ID NO:86, SEQ ID NO:89, SEQ ID NO:120, SEQ ID NO:128, SEQ ID
9 NO:139, SEQ ID NO:153, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:169, SEQ ID
10 NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:183, SEQ ID NO:202, SEQ ID
11 NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID
12 NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID
13 NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID
14 NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID NO:340, SEQ ID NO:351, SEQ ID
15 NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID NO:406, SEQ ID NO:408, SEQ ID
16 NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID NO:437, SEQ ID NO:439, SEQ ID
17 NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID NO:481, SEQ ID NO:484, SEQ ID
18 NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID NO:519, SEQ ID NO:521, and SEQ
19 ID NO:523;

20 (ii) determining the functional effect of the compound upon the nucleic
21 acid or polypeptide; and

22 (iii) expressing the nucleic acid or polypeptide in a cell, contacting the
23 nucleic acid or polypeptide with the compound, and determining the phenotypic or
24 chemical effect upon the cell.

1 25. A method of modulating angiogenesis in a subject, the method
2 comprising the step of administering to the subject a therapeutically effective amount of a
3 compound identified using the method of claim 1.

1 26. The method of claim 25, wherein the subject is a human.

1 27. The method of claim 25, wherein the compound is an antibody.

1 28. The method of claim 25, wherein the compound is an antisense
2 molecule.

THIS PAGE BLANK (USPTO)

12 NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:329, SEQ ID NO:330, SEQ ID
13 NO:340, SEQ ID NO:351, SEQ ID NO:365, SEQ ID NO:377, SEQ ID NO:384, SEQ ID
14 NO:406, SEQ ID NO:408, SEQ ID NO:419, SEQ ID NO:421, SEQ ID NO:428, SEQ ID
15 NO:437, SEQ ID NO:439, SEQ ID NO:445, SEQ ID NO:456, SEQ ID NO:462, SEQ ID
16 NO:481, SEQ ID NO:484, SEQ ID NO:493, SEQ ID NO:496, SEQ ID NO:498, SEQ ID
17 NO:519, SEQ ID NO:521, and SEQ ID NO:523.

1 35. An isolated nucleic acid, wherein the nucleic acid hybridizes
2 under stringent conditions to a polynucleotide comprising a nucleotide sequence selected
3 from the group consisting of SEQ ID NO:165, SEQ ID NO:202, SEQ ID NO:210, SEQ
4 ID NO:218, SEQ ID NO:227, SEQ ID NO:232, SEQ ID NO:248, SEQ ID NO:274, SEQ
5 ID NO:285, SEQ ID NO:286, SEQ ID NO:297, SEQ ID NO:307, SEQ ID NO:308, SEQ
6 ID NO:317, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:323, SEQ ID NO:324, SEQ
7 ID:329, and SEQ ID:330.

1 36. The isolated nucleic acid of claim 35, wherein the nucleic acid
2 comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:165,
3 SEQ ID NO:202, SEQ ID NO:210, SEQ ID NO:218, SEQ ID NO:227, SEQ ID NO:232,
4 SEQ ID NO:248, SEQ ID NO:274, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:297,
5 SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:320,
6 SEQ ID NO:323, SEQ ID NO:324, SEQ ID:329, and SEQ ID:330.

1 37. An isolated polypeptide comprising an amino acid sequence
2 selected from the group consisting of SEQ ID NO:287, SEQ ID NO:298, SEQ ID
3 NO:309, SEQ ID NO:319, SEQ ID NO:325, and SEQ ID NO:331.

THIS PAGE BLANK (USPTO)

INFORMAL SEQUENCE LISTING

SEQ ID NO:1

ABC Transporter

GL2_96_2C01_G3F1

ATGTAGAACATGCGCTGGGGTGGGGGCTTGTAGAGCACACCACCGTACGTGGGCCAGAGCCCACCCAGGATCCGG
AATAGGGAGCTCTTGCCGCGAGCCATTGGGGCCTGTGATGAGCAGATGCATGCCTATCCTCCACCTGATGTTGAGG
CTGGGGCACCACAACCTCTCTGANGGCGTGAACGAATCGGGGAATGTTTCCATCCCGTGCGAAGGAAGTCTNNGTG
TCGACGCGCTCTCGNCTCATCACAACACTCCAACAAACACGTGGGCTCGGAATCTTCAAGGGGGGGGGCCATCCAG
AAAACAGCGAACAACCACTGAACCTGGGGCGTAGATGAGGCTCNCACCCAAGATAAAACACCCCGGGGCGCCTT
GGGAAGACCCGGGTCTCCCGTCCCTTAANACGTTCTCCCCATGTGGCGGGGCCCCACTCCCTTCTTGGTNACAACN
GGCTTGGCACCAAGACCAGGCCGCTCGCAACACCGCAGTCCGCGTTTTTCTCAGACAGCAAAGTTTANCNCCCCCT
TTGTAGGAAGAAAACCCAAAGAGTTTCCGATTCCACAGGAGTTTGGGCCGCAATGCAACAGNCCCAGGAGCGGGG
NNGCCGCATTCGAGTCTAGCGGATTTATATATAGACACGCCGCAACTAGNGAGCGTCCCTCGAAAAGAGCCCN
AGNTTACACNCCCGGGGGTATCCGGTCAAAGCAGCCNTTACCCTCCCTTNTTTTGGTGGTTAATACGGCG
GAAAAAC

SEQ ID NO:2

ABC transporter>

GGACGCGCTGGTGCCCCGGGGAGGGGCGCCACCGGGGGAGGAGGAGGAG
GAGAAGGTGGAGAGGAAGAGACGCCCCCTCTGCCGAGACCTCTCAAGGC
CCTGACCTCAGGGGCCAGGGCACTGACAGGACAGGAGAGCAAGTTCTC
CACTTGGGCTGCCCCAAGAGGGCCGCGACCTTGGAGGGCCCTGAGCCACC
GCACCAGGGGCCCCAGCACCCCCGGGGGCTAAAGCGACAGTCTCAGG
GGCCATCGCAAGGTTTCCAGTTGCCCTAGACAACAGGCCCAGGGTCAGAGC
AACAATCCTTCCAGCCACCTGCCCTCAACTGCTGCCCCAGGCACCAGCCCC
AGTCCCTACGCGGCAGCCAGCCCAGGTGACATGCCGGTGCTCTCCAGGCC
CCGGCCCTGGCGGGGGAACACGCTGAAGCGCACGGCCGTGCTCCTGGCCC
TCGCGGCCTATGGAGCCCAAGTCTACCCCTTGGTGCGCCAGTGCCTG
GCCCCGGCCAGGGGTCTTCAGGCGCCCGCGGGGAGCCACGCAGGAGGC
CTCCGGGGTTCGCGCGGCCAAAGCTGGCATGAACCGGTATTCTCTGCAGC
GGCTCCTGTGGCTCCTGCGGCTGCTGTTCCCCCGGGTCTGTGCGGGAG
ACGGGGCTGCTGGCCCTGCACTCGGCCGCTTGGTGAGCCGCACCTTCCT
GTCCGTGTATGTGGCCCGCCTGGACGGAAGGCTGGCCCGCTGCATCGTCC
GCAAGGACCCGCGGGCTTTTGGCTGGCAGCTGCTGCAGTGGCTCCTCATC
GCCCTCCCTGCTACCTTCGTCAACAGTGCCATCCGTTACCTGGAGGGCCA
ACTGGCCCTGTCGTTCCGCAGCCGTCTGGTGGCCACGCCTACCGCTCT
ACTTCTCCGACGACACTACTACCGGGTCAGCAACATGGACGGGCGGCTT
CGCAACCCTGACAGTCTCTGACGGAGGACGTGGTGGCTTTGCGGCCTC
TGTGGCCACCTCTACTCCAACCTGACCAAGCCACTCTGGACGTGGCTG
TGACTTCCTACACCCTGCTTCGGGCGGCCCGCTCCCGTGGAGCCGGCACA
GCCTGGCCCTCGGCCATCGCCGCGCTCGTGGTGTCTCTCACGGCCAACGT
GCTGCGGGCCTTCTCGCCCAAGTTTCGGGGAGCTGGTGGCAGAGGAGCGC
GGCGGAAGGGGGAGCTGCGCTACATGCACTCGCGTGTGGTGGCCAACCTCG
GAGGAGATCGCCTTCTATGGGGGCCATGAGGTGGAGCTGGCCCTGCTACA
GCGCTCCTACAGGACCTGGCCTCGCAGATCAACCTCATCTTCTGGAAC
GCCTGTGGTATGTTATGCTGGAGCAGTTCCCTCATGAAGTATGTGTGGAGC
GCCTCGGGCCTGCTCATGGTGGCTGTCCCCATCATCACTGCCACTGGCTA
CTCAGAGTCAGATGCAGAGGCGTGAAGAAGGCAGCCTTGGAAAAGAGG
AGGAGGAGCTGGTGAGCGAGCGCACAGAAGCCTTCACTATTGCCCCGAAC
CTCCTGACAGCGGCTGCAGATGCCATTGAGCGGATCATGTGTCGTACAA
GGAGGTGACGGAGCTGGCTGGCTACACAGCCCGGGTGCACGAGATGTTCC
AGGTATTTGAAGATGTTTCAGCGCTGTCACTTCAAGAGGCCAGGGAGCTA
GAGGACGCTCAGGCGGGTCTGGGACCATAGGCCGGTCTGGTGTCCGTGT
TGAGGGCCCCCTGAAGATCCGAGGCCAGGTGGTGGATGTGGAACAGGGGA
TCATCTGCGAGAACATCCCCATCGTCACGCCCTCAGGAGAGGTGGTGGTG
GCCAGCCTCAACATCAGGGTGGAGGAAGGCATGCATCTGCTCATCACAG

CCCCCAATGGCTGCGGCAAGAGCTCCCTGTTCCGGATCCTGGGTGGGCTCT
 GGCCACGTACGGTGGTGTGCTCTACAAGCCCCACCCAGCGCATGTTT
 TACATCCCGCAGAGGCCCTACATGTCTGTGGGCTCCCTGCGTGACCAGGT
 5 GATCTACCCGGACTIONAGTGGAGGACATGCAAAGGAAGGGCTACTCGGAGC
 AGGACCTGGAAGCCATCCTGGACGTCGTGCACCTGCACCACATCCTGCAG
 CGGGAGGGAGGTTGGGAGGCTATGTGTGACTGGAAGGACGTCCTGTGCGG
 TGCGGAGAAGCAGAGAATCGGCATGGCCCGCATGTTCTACCACAGGCCCA
 AGTACGCCCTCCTGGATGAATGCACCAGCGCCGTGAGCATCGACGTGGAA
 10 GGCAAGATCTTCCAGGCGGCCAAGGACGCGGGCATTGCCCTGCTCTCCAT
 CACCCACCGGCCCTCCCTGTGGAAATACCACACACACTTGCTACAGTTTCG
 ATGGGGAGGGCGGCTGGAAGTTGAGAAGCTGGACTCAGCTGCCCGCCTG
 AGCCTGACGGAGGAGAAGCAGCGGCTGGAGCAGCAGCTGGCGGGCATTCC
 CAAGATGCAGCGGCGCCTCCAGGAGCTCTGCCAGATCCTGGGCGAGGCCG
 15 TGCCCCCAGCGCATGTGCCGGCACCTAGCCCGCAAGGCCCTGGTGGCCTC
 CAGGGTGCCTCCACCTGACACAACCGTCCCCGGCCCCCTGCCCGCCCCCA
 AGCTCGGATCACATGAAGGAGACAGCAGCACCCACCCATGCACGCACCCC
 GCCCCTGATGCCTGGCCCCCTCCTCCTAGAAAACCCCTTCCCGCCCTCGGG
 AAAGTAGATGTGGAGGGTGGCGCCCTGCGTAACCCCTCGCCCTGTCCCTCC
 CACTCCCTGGGGCGCTGTTCCACAGTGACTGGGCCCCGTGCCAGGGCAGT
 20 GAGTCCCTCTACTTTGCTCCGTGGAGGAAGCTGGGGTACAAGGGGCCAGT
 GCTGGCCACACAGCAGCGCAGCCGAGCCCCAGGAGCCCGTCAGGCCACAG
 CCCCTGGCACTGCAGGTGGCCTCCCTCCAGAGACTCGAGTCCCCATGATT
 CCCTCCTCGTCAGTCTCTCAAAGACCCCATGGTCCATCCCCTGAGGGTGG
 TCAGCCAAGGCTCCCGTTCCGTGGGATGCCATAAAAGCCGCCAGTGGGA
 25 CCCACAGTCACACAGAGCGCCTCACCTGCATCCTCTCCCCCACAGAGCC
 CCAAAGATCCCACGGGAGAGGGGAGAGGGACGCACAGCACTGCCTGCCAA
 GCGAGAATGCAGGCCCCGCCCCCTCGGCCCTCACCACCTCTTTCTACAG
 CCTAATTTATTGGATTCCCTATTCTGTAGCCATCTCCGTGGCCAATGTGAC
 TACCGTGCCAGCAGCGGGGGCGGCCACGCTCTGAGTCCCGTGGGGCCCC
 30 GGCTCCACCGGTGCCAAACCCAGCCCCCTGCGGCCGTACCCCGCCAGCC
 TACACTGCCAGCCGCCACCGGGGCACACGGGCCTCTGCTTGCCAGCCAGG
 AGTGGCGACACCATGTTCCAGCTCAGTGCCAAAGAGGGGTACCAGGGG
 GAGCTGTCTGCGGAGCCAGCGCCTGCCCCGAGAGAGACCCACCGCCACCG
 35 TGTGCCCTTTCCCGGGCCCTCAGCCCTCGGGCCGGGCACCACCCCACTCC
 CCCCAGTAAAAGCCTCCACTGGCAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO:3

>gi|7262392|ref|NM_000033.2| Homo sapiens ATP-binding cassette, sub-family
 D (ALD), member 1 (ABCD1), mRNA
 40 GCGGACGGACGCGCCTGGTGCCCCGGGGAGGGGCGCCACCGGGGAGGAGGAGGAGAAGGTGGAGAG
 GAAGAGACGCCCCCTCTGCCCGAGACCTCTCAAGGCCCTGACCTCAGGGGCCAGGGCACTGACAGGACAG
 GAGAGCCAAGTTCTTCACTTGGGCTGCCCCGAAGAGGCCGCGACCCCTGGAGGGCCCTGAGCCCACCGCAC
 CAGGGGCCCCAGCACCCCGGGGGCCTAAAGCGACAGTCTCAGGGGCCATCGCAAGGTTTCCAGTTGC
 45 CTAGACAACAGGCCCAGGGTCAGAGCAACAATCCTTCCAGCCACCTGCCTCAACTGCTGCCCCAGGCACC
 AGCCCCAGTCCCTACGCGGACGACCCAGGTCAGATGCCGGTGTCTCCAGGCCCGGCCCTGGCGGG
 GGAACACGCTGAAGCGCACGGCCGTGCTCCTGGCCCTCGCGGCTATGGAGCCACAAAGTCTACCCCTT
 GGTGCGCCAGTGCCTGGCCCCCGGCCAGGGGTCTTCAGGCGCCCGCGGGGAGCCACGCAGGAGGCCTCC
 GGGGTGCGGGCGGCCAAAGCTGGCATGAACCGGGTATTCTTGCAGCGGCTCCTGTGGCTCCTGCGGCTGC
 50 TGTCCCCCGGGTCTGTGCGGGGAGACGGGGCTGCTGGCCCTGCACTCGGCCGCCTTGGTGAGCCGCAC
 CTTCCTGTGCGGTGTATGTGGCCCGCCTGGACGGAAGGCTGGCCCCGTGCATCGTCCGCAAGGACCCGCGG
 GCTTTTGGCTGGCAGCTGCTGCAGTGGCTCCTCATCGCCCTCCCTGCTACCTTCGTCAACAGTGCCATCC
 GTTACCTGGAGGGCCAACCTGGCCCTGTGCTTCCGCGAGCCGTCTGGTGGGCCACGCCTACCGCCTCTACTT
 CTCCCAGCAGACCTACTACCGGGTCAGCAACATGGACGGGCGGCTTCGCAACCCCTGACCAGTCTCTGACG
 GAGGACGTGGTGGCCTTTGCGGCCCTCTGTGGCCACCTCTACTCCAACCTGACCAAGCCACTCCTGGACG
 55 TGGCTGTGACTTCTACACCTGCTTCGGGCGGCCGCTCCCGTGGAGCCGGCACAGCCTGGCCCTCGGC
 CATCGCGGCTCGTGGTGTTCCTCACGCCCAACGTGCTGCGGGCCTTCTCGCCCAAGTTTCGGGGAGCTG
 GTGGCAGAGGAGGCGCGGCGGAAGGGGAGCTGCGCTACATGCACTCGCGTGTGGTGGCCAACCTCGGAGG
 AGATCGCCTTCTATGGGGGCCATGAGGTGGAGCTGGCCCTGCTACAGCGCTCCTACCAGGACCTGGCCTC
 60 GCAGATCAACCTCATCCTTCTGGAACGCCTGTGGTATGTTATGCTGGAGCAGTTCTCATGAAGTATGTG
 TGGAGCGCCTCGGGCCTGCTCATGGTGGCTGTCCCCATCATCACTGCCACTGGCTACTCAGAGTCAGATG
 CAGAGGCCGTGAAGAAGGCAGCCTTGGAAAAGAAGGAGGAGGAGCTGGTGAGCGAGCGCACAGAAGCCTT

CACTATTGCCCGCAACCTCCTGACAGCGGCTGCAGATGCCATTGAGCGGATCATGTCGTCGTACAAGGAG
 GTGACGGAGCTGGCTGGCTACACAGCCCGGGTGCACGAGATGTTCCAGGTATTTGAAGATGTTTCAGCGCT
 GTCACTTCAAGAGGCCCAGGGAGCTAGAGGACGCTCAGGCGGGGTCTGGGACCATAGGCCGGTCTGGTGT
 5 CCGTGTGGAGGGCCCCCTGAAGATCCGAGGCCAGGTGGTGGATGTGGAACAGGGGATCATCTGCGAGAAC
 ATCCCCATCGTCACGCCCTCAGGAGAGGTGGTGGTGGCCAGCCTCAACATCAGGGTGGAGGAAGGCATGC
 ATCTGCTCATCACAGGCCCAATGGCTGCGGCAAGAGCTCCCTGTTCCGGATCCTGGGTGGGCTCTGGCC
 CACGTACGGTGGTGTGCTCTACAAGCCCCACCCAGCGCATGTTCTACATCCCGCAGAGGCCCTACATG
 TCTGTGGGCTCCCTGCGTGACAGGTGATCTACCCGGA CTAGTGGAGGACATGCAAAGGAAGGGCTACT
 CGGAGCAGGACTGGAAGCCATCCTGGACGTGCTGCACCTGCACCACATCCTGCAGCGGGAGGGAGGTTG
 10 GGAGGCTATGTGTGACTTGAAGGACGTCTGTGCGGTGGCGAGAAGCAGAGAATCGGCATGGCCCCGCATG
 TTCTACCACAGGCCCAAGTACGCCCTCCTGGATGAATGCACACAGCGCCGTGAGCATCGAGTGGGAAGCA
 AGATCTTCCAGGCGGCCAAGGACGCGGGCATTGCCCTGCTCTCCATCACCCACCGCCCTCCCTGTGGAA
 ATACCACACACACTTGCTACAGTTCGATGGGAGGGCGGCTGGAAAGTTCGAGAAGCTGGACTCAGCTGCC
 CGCCTGAGCCTGACGGAGGAGAAGCAGCGGCTGGAGCAGCAGCTGGCGGGCATTCCTAAGATGCAGCGGC
 15 GCCTCCAGGAGCTCTGCCAGATCCTGGGCGAGGCCGTGGCCCCAGCGCATGTCCCGGCACCTAGCCCGCA
 AGGCCCTGGTGGCCTCCAGGGTGCCCTCCACCTGACACAACCGTCCCCGGCCCCCTGCCCGCCCCCAAGCT
 CGGATCACATGAAGGAGACAGCAGCACCCACCCATGCACGCACCCCCGCCCTGCATGCCTGGCCCCCTCCT
 CCTAGAAAACCTTCCCGCCCTCGGGAAGTAGATGTGGAGGGTGGCGCCCTGCGTAACCTCGCCCTGT
 CCCTCCCACTCCCTGGGGCGCTGTTCCACAGTGACTGGGCCCTGTCCAGGGCAGTGAGTCCCTCTACTTT
 20 GCTCCGCGGAGGAAGCTGGGGTACAAAGGGGCCAGTGCTGGCCACACAGCAGCGCAGCCGAGCCCCAGGA
 GCCCCGTCAGGCCACAGCCCCCTGGCACTGCAGGTGGCCTCCCTCCAGAGACTCGAGTCCCCATGATTCCCT
 CCTCGTCAGTCTCTCAAAGACCCCATGGTCCATCCCTGAGGGTGGTCAGCCAAGGCTCCCGTTCCGTGG
 GATGCCATAAAAGCCGCCCAGTGGGACCCACAGTCACACAGAGCGCCTCACCTGCATCCTCTCCCCACA
 AGAGCCCCAAAGATCCCACGGGAGAGGGGAGAGGGACGCACAGCACTGCCTGCCAAGCGAGAATGCAGGC
 25 CCGCCCCCTCGGCCCTCACCACCTCTTTCTACAGCCTAATTTATTGGATTCCCTATTCTGTAGCCATCT
 CCGTGGCCAATGTGACTACCGTGCCAGCAGCGGGGCGGCCAGCCTCTGAGTCCCGTGGGGCCCCGGCT
 CCCACCGGTGCCAAACCCAGCCCCCTGCGGCCGTACCCCGCCAGCCTACACTGCCAGCCGCCACCGGGGC
 ACACGGGCCTCTGCTTGCCAGCCAGGAGTGCGGACACCATGTTCCAGCTCAGTGCCAAAGAGGGGTAC
 30 CAGGGGAGCTGTCTGCGGAGCCAGCGCCTGCCCGAGAGAGACCCACCGCCACCGTGTGCCTTTCCCGG
 GCCCTCAGCCCTCGGGCCGGGCACCACCCCAAGTCCCCCAGTAA

SEQ ID NO:4

>gi|7262393|ref|NP_000024.2| ATP-binding cassette, sub-family D (ALD),
 member 1; adrenoleukodystrophy protein [Homo sapiens]
 35 MPVLSRPRPWRGNTLKR TAVLLALAA YGAHKVYPLVRQCLAPARGLQAPAGEPTQEASGVAAAKAGMNRV
 FLQRLWLLRLLFPRLVCRETG LLLALHSAALVSR TFLSVYVARLDGRLARCIVRKDPRAFGWQLLQWLLI
 ALPATFVNSAIRYLEGQLALSFRSRLVAHAYRLYFSQQTYYRVSNMDGRLRNP DQSLTEDVVAFAASVAH
 LYSNLTKPLLDVAVTSYTL LRAARSRGAGTAWPSA IAGLVVFLTANVLRAFSPKFGELVAEEARRKGELR
 40 YMHSRVVANSEEIAFYGGHEVELALLQRSYQDLASQINLILLERLWYVMLEQFLMKYVWSASGLLMVAVP
 IITATGYSESDAEAVKKAAL EKKKEELVSRTEAFTIARNLLTAAADAIERIMSSYKEVTELAGYTARVH
 EMFQVFEDVQRCHFKRPRELEDAQAGSGTIGRSGVRVEGPLKIRGQVVDVEQGIICENIPIVTPSGEVVV
 ASLNIRVEEGMHLLITGPN GCGKSSLFRILGGLWPTYGGVLYKPPPQRMFYIPQRPYMSVGLRDQVIYP
 DSVEDMQRKGYSEQDLEAILD VVHLHHILQREGWEAMCDWKDVLSGGEKQIRIGMARMFYHRPKYALLDE
 45 CTSAVSIDVEGKIFQAAKDAGIALLSITHRPSLWKYHTLLQFDGEGGWKFELDSAARLSLTBEKQRL
 QQLAGIPKMQRRLQELCQILGEAVAPAHVPAPSPQGPGLQAST

SEQ ID NO:5

PCBP1

GL2_131_2_M13F

50 GCTGCGGATCAACATCTCGGAGGGGAATTGTCCGGAGAGAATCATCACATCATGACCGGCCCAACATGCCATC
 TTAAAGAGCTTTTCGTATGATCATCGACAAGACTGGAGGAAGATATCAACAGCTCCATGACCAACAGTACCGCGG
 CACAGCAGGCCCCACGGGTACCCCTGAGGCTGGATGGTGCCGGCCACCCAGATGCGGCTCCCTGATANGGAAAAG
 GCGGGTGTAAAGATCAAAGAGATCANNAGAGAGTACGGGGNGCGCAGGTCCAAGGAATGGCAGGAGGGATATGCAT
 55 GCCCAAAAACAACCGAGACGGGCCAAANACCATAGCTGGACAGTGCCGAAGGACTGTACCNGAAGAT

SEQ ID NO:6

GL2-77-2M13R

60 CGCTGCGGCCTCCCGCCGCTCCCGCTCGCTCCCGCGGCCTCGCTCGCCTCGCGCCGGCAGTTTTGGGCCTACAC
 CTCCCCCTCCCCCGCCAGCCGCCAAAGACTTGACCACGTAACGAGCCCAACTCCCCGAACGCCGCCCGCGCTCG
 CCATGGATGCCGGTGTGACTGAAAGTGGACTAAATGTGACTCTCACCATTCCGGCTTCTTATGCACGGAAAGGAAG

TAGGAAGCATCATTTGGGAAGAAAGGGGAGTCCGTATAAGAGGATCCGCGAGGAGAGTGGNGCGCGGATCAACATC
TCGGAGGGGAATTTGTACCGGAGAGAATCATCACTCTGACCGGCNCCCAACAAATGCACAANCNTAAAGGACNTA
AACGACAAATAGAAATCATACGACAAANCTAGAGAGAAGATATCAACAGACTCCAATGAACAAAANAAACCGCAGA
GACAGACAGGAACCCCNNGGAAC

SEQ ID NO:7

GL2_209_3_M13R

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
ACCGATCTTGCCCGCCCGCGCAGATGACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCC
ACTTGGCGGAGACCTGGGCAGAGGCCGGTTCTCCAGGGCAATGACCCAGGCAAATCTGCTTGAGACACACCTA
CGGTGCACAGACTGGCGGGGCAACGGCACAGAGCAGAAGTTGCGGTGAACAGCTGGGCACACAAGGACGA

SEQ ID NO:8

GL2_131_3_M13F

CATGGATGCCGGTGTGACTGAAAGTGGACTAAATGTGACTCTCACCATTTCGGCTTCTTATGCACGGAAGGAAGT
AGGAAGCATCATTTGGGAAGAAAGGGGAGTCCGTATAAGAGGATCCGCGAGGAGAGTGGCGCGCGGATCAACATCTC
GGAGGGGAATTTGTCCGGAGAGAATCATCACTCTGACCGGCCCCACCAATGCCATCTTTAAGACTTTTCGCTATGAT
CATCGACAAGCTGGAGGAAGATATCAACAGCTCCATGACCAACAGTACCGCGGCCAGCAGGCCCCCGGTACCCCT
GAGGCTGGTGGTGCCGGCCACCCAGTGCGGCTCCCTGATTGGGAAAGGCGGGTGTAAAGATCAAAGAGATCCGCG

SEQ ID NO:9

GL2_205_C05_G3F1

AGAGGCGCAGCGGTTCGACGCGNTGGCTCAGTGTGCTGGGTTCTTGCGGCTTGAATCGAGTAGGCATCCAGAGGTG
GTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGCACCGATCTTGGCGCCCGCGCAGATGAC
TGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCCCTTGCGGAGACTGGGAGAGCGTCTCCAG
CATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCACGCCAGCGATGGTGTATGGCCCGCTCGGT
GGAGTTTGGGCAGCATATCCCCGCAACTGGACCTGCGCCCCCGTACTCTCGCGGATCTCTTTGATCTTACACCC
GCCTTTCCCAATCAGTGGAGCCGCACTGGGTGGGCGGCAACAACAGCTCAGGGTGACCGGGGGCCTGCTGGCCGC
GGTACTGTTTGGTCATGGAGCTGTTGATATCTTCCCTCCAGCTTGTGATGATCATAGCGAAAGCCTTAAAGATGG
CATTGGTGGGGCCGGTCAGAGTGATGATTCTCTCCGACAATTTCCCTCCGAGATGTTTGATCCGGGCGCCACTC
TCCTCGCGGATCCTCTTAACGAATCCCCCTTCTTCCCAATGATGCTTCTACTTCTTCTTCCGTGCATAGGAAGCCG
AATGGTGAGAGTCACATTTTAGTCCACTTTTCAGTCACACCGGGGATCCATGGNCGANCGGGCGGGGACGGGCGT
ACGGGGGAATTGGGCTCGTAACAGTTGGNCCAGNTCNTTAGGCGGTGAGAGGGAGGGAAGGGGAGGCCAACAAATGG
CGGGGAGAAGAAGGCACGACAACAGAGTGCGGAGGACAACCAACACAACAAGGTAAAAAA

SEQ ID NO:10

GL2_131_3_M13R

CTTGAATCGAGTAGGCATCCAGAGNGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTG
CACCGATCTTGGCCCGCCCGCGTAGATGACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTC
CCTTGCGGAGACTGGGAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGC
ACGCCAGCGATGGTGTATGGCCCGCTCGGNGAGAGTTAGGGCAGCAATATCCCCACGAACAACTGGAAC TGACG
GCCCCCGAAACTTACGCGNAGATACTCNTTAAGAACTTNACACCCGGAATATACCCAAATCAGAGAAGCCGCA
ACATGGGTGGCACCGAAAAACAACAGAACTCCAGGGGATGAACCGGAGGGGACATAGAAAGGGAACAGACGGGGA
TACCATAGGTCTGGCGGTCCCAATGGGGAGGCTGCNTATGAATAAATCCTTACCCTTCCCAANCACATAGGNA
TCCGGAATGGAAATCAAATAGGCGGAAAAANGTCCCTTTAAACAGAAAAATNGGGCAACTTNGNAAGGGGGGGAC
CACGAGACANAACAAACCGAAGACATACGCAACCAGGGAACAAAANTTACCCGACACGCAGAGAAGGNTGGACA
NCACGGCACAGACCACAGTACCCATCNAGACGGACCCG

SEQ ID NO:11

GL2_209_1_M13R

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
ACCGATCTTGGCCCGCCCGCGCAGATGACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCC
CTTGCGGAGACTGGGAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCA
CGCCAGCGATGGTGTATGGCCCGCTCGGTGGAGTTGGGCAGCATATCCCCCGCACACCTGGACCTGCGCCCCGTA
CTCTCGCGGATCTCTTATGATCTTAACAACACACGACACNTTTTTNCANCAAAATCCAGGGAGCCGACAACATG
GAGTGAGCCCGGACACCACACAGACCTCAGGGTGACCGGGGGCTGCATGGCCGACGGATACNGAANGGTATGG
ACGCTGTGATATCTTCTCCAGCTAGTCGATGATCATAGCGAAAAGCCCTTTCACAACGAAATGAGGCCATTNG
AGATNGGGGCGGCCCGGGATCCAGGAAGGTTGCGACTAGGAACCTNACGTACCATTATCCCCAGGGAGAAAACNCC
AGAACAATATTACCCACCCACCGATTACCCCAAGAAAGAA

SEQ ID NO:12

GL2_131_2_M13R

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
GCCGATCTTGCCCGCCCGCGCAGATGACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCC
5 CTTGCGGAGACTGGGAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCA
CGCCAGCGATGGTGATGGCCCGCTCGGTGGAGTTGGGCAGCATATCCCCACGAGCANCTGGACCTGCGCCCCCGA
TACTCTCGACGGATCTCTTTGATCTTACACCCGACCTTTCCCAATCAGGGAAGCCGCACATGGGATGGCAGGCA
CCACCAGACCTCAGGGATGACCAGGGAGGCCTGCATGGACCGCAGGATACTGATAGGATCATGGAGNCTGATAGA
10 TATCTTCCCTCCAGACTATGNTCGATGATCAATAAGGCGGAAAAGGCCATAAANAGNAATGGGCAATTAGGGATC
GGGAGGACCAGGATCAAGA

SEQ ID NO:13

GL2_81_1_M13F

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCAGTAATGGCCGGGTTTTTTTGGGTNNGGGGCCAAAT
15 TTTCCGGGGGAAGGAGTAAAGGCCCCACCGGGACCAATTGGCAGGGTCCGCTGCNACCGATACTTGGCCTGCCGC
GCAGATGACTGGGGAGCTGGTCCGCATGGGCTGGTACGGAATGGTCATGACTCTCCCTTGCGGAGACTGGGAGAG
CGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCACGCCAGCGATGGTGATGGC
CCGCTCGGTGGAGTTGGGCAGCATATCCCCGCCACCTGGACCTGCGCCCCCGTACTCTCGCGGATCTCTTTGAT
20 CTTACACCCGCTTTCCCAATCAGGGAGCCGCACTGGGTGGCCGGCACCACCAGCCTCAGGGTGACCGGGGGCCT
GCTGGCCCGGTACTGTTGGTTCATGGGGCTGTTGATATCTTCCCTCCAGCTTGTCTG

SEQ ID NO:14

GL2_77_1_M13F

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
25 ACCGATCTTGCCCGCCCGCGCAGATACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCCC
TTGCGGAGACTGGGAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACCCCTCGGTGACAGACTGCGGCAC
GCCAGCGATGGTGATGGCCCGCTCGGTGGAGTTGGGCAGCATATCCCCGCCACCTGGACCTGCGCCCCCGTACT
CTCGCGGATCTCTTTGATCTTACACCCGCTTTCCCAATCAGGGAGCCGCACTGGGTGGCCGGCACCACCAGCCT
30 CAGGGTGACCGGGGGCCTGCTGGCCGCGGTACTGTTGGTTCATGGAGCTGTTGATAT

SEQ ID NO:15

GL2_77_1_M13R

AGCTTGGTACGAGCTCGGATCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTCGAGGAGAGTGCGCGCGC
GGATCAACATCTCGGAGGGGAATTGTCCGAGAGAATCATCACTCTGACCGGCCCCACCAATGCCATCTTTAAGG
35 CTTTCGCTATGATCATCGACAAGCTGGAGGAAGATATCAACAGCTCCATGACCAACAGTACCGCGGCCAGCAGGC
CCCCGGTCAACCTGAGGCTGGTGGTGGCGGCCACCCAGTGCAGCTCCCTGATTGGGAAAAGGCGGTGTAAGATCA
AAGAGATCCGCGAGAGTACGGGGGCGCAGGTCCAGGTGGCGGGGGATATGCTGCCCAACTCCACCGAGCGGGCCA
TCACCATCGCTGGCGTGCCGCACTCTGTACCGAGGGTGTCAAGCAGATTTGCTTGGTTCATGCTGGAGACGCTCT
40 CCCAGTCTCCGCAAGGGAGAGTCATGACCATTCCTGACCAGCCATGCCGGCCAGCTCCCCAGTATCTGCGCGGG
CGGCCAAGATCGGTGCAGCGACGCTGCGGGCTACCCCATGCCACCCATGACCTGGAGGGACCACCTCTGGATGC
CTACTCGATTCAAG

SEQ ID NO:16

GL2_77_2_M13F

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
45 ACCGATCTTGCCCGCCCGCGCAGATACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCCC
TTGCGGAGACTGGGAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCAC
GCCAGCGATGGTGATGGCCCGCTCGGTGGAGTTGGGCAGCATATCCCCGCCACCTGGACCTGCGCCCCCGTACT
CTCGCGGATCTCTTTGATCTTACACCCGCTTTCCCAATCAGGGAGCCGCACTGGGTGGCCGGCACCACCAGCCT
50 CAGGGTGACCGGGGGCCTGCTGGCCGCGTACTGTTGGTTCATGGAGCTGTTGATATCTTCCCTCCAGCTTGTGAT
GATCATAGCGAAAGCCTTAAAGATGGCATTGGTGGGGCCGGTCAGAGTGATGATTCTCTCCGGACAATTCCTCCTC
CGAGATGTTGATCCGCGCGCCATCTCTCGCGGATCTCTTAAACCGATCCCCTTTCTTCCCAATGATGCTTCTA
CTTCTTTCCGTGCATAGAA

SEQ ID NO:17

GL2_77_3_M13F

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
55 ACCGATCTTGCCCGCCCGCGCAGATACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCCC
TTGCGGAGACTGGGAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCAC
GCCAGCGATGGTGATGGCCCGCTCGGTGGAGTTGGGCAGCATATCCCCGCCACCTGGACCTGCGCCCCCGTACT

CTCGCGGATCTCTTTGATCTTACACCCGCCCTTCCCAATCAGGGAGCCGCACTGGGTGGCCGGCACCACCAGCCT
CAGGGTG

SEQ ID NO:18

GL2_77_4_M13F

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
ACCGATCTTGGCCGCCCCGCGCAGATACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCCC
TTGCGGAGACTGGGAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCAC
GCCAGCGATGGTGATGGCCCCGCTCGGTGGAGTTGGGCAGCATATCCCCCGCCACCTGGACCTGCGCCCCCGTACT
CTCGCGGATCTCTTTGATCTTACACCCGCCCTTCCCAATCAGGGAGCCGCACTGGGTGGCCGGCACCACCAGCCT
CAGGGTGACCGGGGGCCTGCTGGCCGCGGTACTGTTGGTCATGGAGCTGTTGATATCTTCTCCAGCTTGTCGAT
GATCATAGCGAAAGCCTTAAAGATGGCATTGGTGGGGCCGGTCAGAGTGATGATTCTCTCCGACAATTCCCCTC
CGAGATGTTGATCCGCGCGCCACTCTCTCGCGGATCCTCT

SEQ ID NO:19

GL2_77_5_M13F

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
ACCGATCTTGGCCGCCCCGCGCAGATACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCCC
TTGCGGAGACTGGGAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCAC
GCCAGCGATGGTGATGGCCCCGCTCGGTGGAGTTGGGCAGCATATCCCCCGCCACCTGGACCTGCGCCCCCGTACT
CTCGCGGATCTCTTTGATCTTACACCCGCCCTTCCCAATCAGGGAGCCGCACTGGGTGGCCGGCACCACCAGCCT
CAGGGTGACCGGGGGCCTGCTGGCCGCGGTACTGTTGGTCATGGAGCTGTTGATATCTTCTCCAGCTTGTCGAT
GATCATAGCGAAAGCCTTAAAGATGGCATTGGTGGGGCCGGTCAGAGTGATGATTCTCTCCGACAATTCCCCTC
CGAGATGTTGATCCGCGCGCCACTCTCTCGCGGATCCTCTTAACCGACTCCCCTTTCTTCCCAATGATGCTTC
CTACTTCCCTTTTCCGTGCATAAGAAAGCCGAAATGGTGAGAG

SEQ ID NO:20

GL2_131_1_M13F

NGCTGCGGATCAACATCTCGGAGGGGAATTGTCCGGAGAGAATCATCACATCATGACCGGCCCCACCAATGCCA
TCTTTAAGAGCTTTCGTATGATCATCGACAAGACTGGAGGAAGATATCAACAGCTCCATGACCAACAGTACCGC
GGCACAGCAGGCCCCACGGGTACCCCTGAGGCTGGATGGTGGCCGCCACCCAGATGCGGCTCCCTGATANGGGAA
AGGCGGGTGTAAGATCAAAGAGATCANNAGAGAGTACGGGGNGCGCAGGTCCAAGGAATGGCAGGAGGGATATGC
ATGCCCCAAAACAACCGAGACGGGCCAAANACCATAGCTGGACAGTGCCGAAGGACTGTCACCNAGAAGAT

SEQ ID NO:21

GL2_131_1_M13R

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
ACCGATCTTGGCCGCCCCGCGCAGATGACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCC
CTTGCGGGGAGAACTGGGGAAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGC
GGCACGCCAGCAATGGTGATGGCCCGCTCGGTGGAGTTGGGCAGCATATCCCCCGCCACCTGGACCTGCGCCCCC
GTACTCTCGCGGATCTCTT

SEQ ID NO:22

GL2_209_2_M13R

GGCGCCAGTGCTGTCNATTCGNTATCTCTCGGCATGGACGAGCTGGTACAAAAAGGAGGAGGGCCCGCAAGTC
GGTGGCAGCGGTGGCTCCAGTGTTGGCTGGGGGGTTCTGCGGCTTGAATCGGAAGGTTTAAGGGGGCATCCAGA
GGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGCACCGATCTTGGCGCCGCGCCAGA
TGACTGGGGGAAGCTGGCCGGCATGGGCTTGGTACGGAA

SEQ ID NO:23

GL2_209_3_M13F

CGCTGCGGATCAACATCTCGGAGGGGAATTGTCCGGAGAGAATCATCACTCTGACCGGCCCCACCAATGCCATCT
TTAAGGCTTTTCGTATGATCATCGACAAGCTGGAGGAAGATATCAACAGCTCCATGACCAACAGTACCGCGGCCA
GCAGGCCCCCGGTACCCCTGAGGCTGGTGGTGCCGGCCACCCAGTGCGGCTCCCTGATTGGGAAAGGCGGGTGTA
AGATCAAAGAGATCCGCGAGAGTACGGGGGCGCAGGTCCAGGTGGCGGGGGATATGCTGCCCCAACTCCACCGAGC
GGCCCATCACCATCGCTGGCGTGGCAGTCTGTGTCACCGAGTGTGTCAAGCAGATTTGCTGGTTCATGCTGGAGA
CGCTCTCCAGTCTCCGCAAGGGAGAGTCATGACCATTCGTAACCCAGCCCCATGCCGGCCAGCTCCCCAGTCATCT
GCGCGGGCGGCCAAAGATCGGNGCAGCGACGCTGCGGGCTACCCACAATAGACACACCATGACCTGGAGGGAAC
AACTACTGGATGCTACATCCGANNTCAAGCCGGAAAAANCCAACAACTGGGAACACACGATAGACACACGTCG
CCTTGGCAGGCATCACACCTATGAA

SEQ ID NO:24

GL2_209_4_M13R

5 CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
ACCGATCTTGGCCGCCCCGCGCAGATGACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCC
CTTGCGGAGACTGGGAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCA
CGCCAGCGATGGTGATGGCCCGCTCGGTGGAGTTGGGCAGCATATCCCCCGCCACCTGGACCTGCGCCCCCGTAC
TCTCGCGGATCTCTTTGATCTTACACCCGCCTTTCCCAATCAGGGAGCCGCACTGGGTGGCCGGCACCACCAGCC
10 TCAGGGTGACCGGGGGCCTGCTGGCCGCGGTACTGTTGGTCATGGAGCTGTTGATATCTTCCCTCCAGACTTGTCTG
ATGATCATAGCGAAAGCCTTAAAGATGGCATTGGTGGGGGCCGGTCAGAGTGATGATTCTCTCCGGACAATTCC
CCTTCGAAGATGTGATCCGCGCGCACATCTCCTCGCGGATCTCTTAAACCGAGTCCCNNTTTCATCCAAGAATGNA
CTCCTACTCCCTATACCGGCCATAANAANGCCGAAATGGGTGAAGAANGTCACACTTTAAGGTCCACACTTATCA
GCACANCCCCGGATCCACAGGGGCAAACGGCCGGGAGCAGGAAGGCCCGAGAGGCCCGGAAAAACCCANACAACA
GGGTTAAATAAAAAATAAAGGGCCGCAAAAAAATCGAAGGAATCGAGAAAAACANGCCGCCAATCGGCACACTAA
15 TGAACGATAACAACATAAAGACAGA

SEQ ID NO:25

GL2_131_4_M13F

20 CGCTGCGGGCTCCCGGCCCCGGCTCGCCATGGATGGCCGGTGTGACTGAAAGTGGACTAAATGTGANTCTCACCAT
TCGGCTTCTTATGCACGGAAGGAAGTAGGAAGCATCATTGGGAAGAAAGGGGAGTCNGTTAAGANGATCCGTGA
GGAGAGTGGCGCGCGGATCAACTATCTCGTGAGGGAATTGTCNCGGAGAGAANCATCACTCTGACCGGCCCCACC
AATGANCANCTTTAAGGCTTGCGCATATGATCATCGACAAGCNGGAGGAAGATATCAACAGANTACCATGACC
CAGTACCGCGGGACAGACANGCCCCCGGNCAGACANGAGGCTGGAGGNAGCCCGACCAGC

25 SEQ ID NO:26

GL2_205_C04_G3F1

30 CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
ACCGATCTTGGCCGCCCCGCGCAGATGACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCC
CTTGCGGAGACTGGGAGAGCGTCTCCAGCATGACCAGGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCA
CGCCAGCGATGGTGATGGCCCGCTCGGTGGAGTTGGGCAGCATATCCCCCGCACTGGACCTGCGCCCCCGTACTC
TCGCGGATCTCTTTGATCTTACACCCGCCTTTCCCAATCAGGGAGCCGCACTGGGTGTGCCGGCACCACCAGCCT
CAGGGTGACCGGGGGCCTGCTGGCCGCGGTACTGTTGGTCATGGAGCTGTTGATATCTTCCCTCCAGCTTGTCTGAT
GATCATAGCGAAAGCCTTAAAGATGGCATTGGTGGGGCCGGTCAGAGTGATGATTCTCTCCGGACAATCCCCCTC
CGAGATGTTGATCCGCGCGCCACTCTCCTCGCGGATCTCTTAAACGATCCCCCTTCTTCCCAATGATGCTTCCA
35 TACTTCTTTCCGTGCTAAGAAGCCGAAATGGTGAGAGTCACCATTTAGGTCCACTTTTCAGTCAACACCGGATC
CATGGGCGAGCGGGGGCAGGACGTACCGGGGNGAGATTGGGCTCGAACAGTGGGCAAAGACAGGACAGA

SEQ ID NO:27

GL2_77_4_M13R

40 CGCTGCGGCTCCCGCCCCGCTCCCGCTCGCTCCCGCGGCCCTCGCTCGCCTCGCGCCGGCAGTTTTGGGCCTACA
CCTCCCCCTCCCCCGCCAGCCGCCAAAGACTTGACCACGTAACGAGCCCAACTCCCCCGAACGCCGCCCGCGCT
CGCCATGGATGCCGGTGTGACTGAAAGTGGACTAAATGTGACTCTCACCATTGGCTTCTTATGCACGGAAGGA
AGTAGGAAGCATCATTGGGAAGAAAGGGGAGTCGGTTAAGAGGATCCGCGAGGAGAGTGGCGCGCGGATCAACAT
CTCGGAGGGGAATTTGTCCGGAGAGAATCATCACTCTGACCGGCCCCACCAATGCCATCTTTAAGGCTTTCGCTA
45 TGATCATCGACAAGCTGGAGGAAGATATCAACAGCTCCATGACCAACAGTACCGCGGCCAGAGCCCCCGGTCA
CCCTGAGGCTGGTNGCCGGCCACCCAGTGCGGCTCCCTGATTGGGAAAGGCGGGTGTAAGATCAAAGAGATC
CGCGAGGTACGGGGGCGCAGGTCCAGGT

SEQ ID NO:28

GL2_77_5_M13R

50 CGCTGCGGCTCCCGCCCCGCTCCCGCTCGCTCCCGCGGCCCTCGCTCGCCTCGCGCCGGCAGTTTTGGGCCTACA
CCTCCCCCTCCCCCGCCAGCCGCCAAAGACTTGACCACGTAACGAGCCCAACTCCCCCGAACGCCGCCCGCGCT
CGCCATGGATGCCGGTGTGACTGAAAGTGGACTAAATGTGACTCTCACCATTGGCTTCTTATGCACGGAAGGA
AGTAGGAAGCATCATTGGGAAGAAAGGGGAGTCGGTTAAGAGGATCCGCGAGGAGAGTGGCGCGCGGATCAACAT
55 CTCGGAGGGGAATTTGTCCGGAGAGAATCATCACTCTGACCGGCCCCACCAATGCCATCTTTAAGGCTTTCGCTA
GATCATCGACAAGCTGGAGGAAGATATCAACAGTCCATGACCAACAGTACCGCGGCCAGAGCCCCCGGTCA
CCTGAGGCTGGTGGTGCCGGCCACCCAGTGCGGCTCCCTGATTGGGAAAGGCGGGTGTAAGATCAAAGAGATCCG
CGAGAGTACGGGGGCGCAGGTCCAGGG

SEQ ID NO:29

GL2_209_4_M13F

CGCTGCGGCCCTCCCGCCCGCGCTCGCCATGGATGCCGGTGTGACTGAAAGTGGACTAAATGTGACTCTCACCAT
 TCGGCTTCTTATGCACGGAAGGAAGTAGGAAGCATCATTTGGGAAGAAAGGGGAGTCGGTTAAGAGGATCCGCGA
 5 GGAGAGTGGCGCGCGGATCAACATCTCGGAGGGGAATTGTCCGGAGAGAATCATCACTCTGACCGGCCCCACCAA
 TGCCATCTTTAAGGCTTTTCGCTATGATCATCGACAAGCTGGAGGAAGATATCAACAGCTCCATGACCAACAGTAC
 CGCGGCCAGCAGGCCCCCGGTCACCTGAGGCTGGTGGTGCCGGCCACCCAGTGGGCTCCCTGATTGGGAAAGG
 CGGGTGTAAAGATCAAAGAGATCCGCGAGAGTACGGGGGCGCAGGTCCAGGTGGCGGGGGATATGCTGCCAACTCC
 10 ACCGAGCGGGCCATCACCATCGCTGGCGTGCCGAGTCTGTTACCCGAAGTGTGTACAGCNAGATTTGCTTGGT
 CATGCTTGGAAAACGGCTTCTCCCCAANTACCTTCCNGCAAAGGGGAGAAGTCCATTGAACCCANTTCCCCCGCN
 AACCAACAGCCCCCAAATGGGCCCCGGGGCCACAGGCTCCCCCAAGGACANATCNGGAGGCCGGGGGACGGGGCC
 CAACGAATCCGGGAGGCAAGACGAANACATGCAGGGCATAAACCCCCCANGGCACACCCATGAACCTGGGAAGGG
 GACCACCTCTGGGAATGGCNAATCGAGTCAAGCCAGAAAAACCAGCACACG

SEQ ID NO:30

GL2_205_C06_G3F1

CTTGAATCGAGTAGGCATCCAGAGGTGGTCCCTCCAGGTCATGGGTGGCATGGGGGTAGCCCGCAGCGTCGCTGC
 ACCGATCTTGGCGCCGCGCAGATGACTGGGGAGCTGGCCGGCATGGGCTGGTACGGAATGGTCATGACTCTCCCT
 TGGCGAGACTGGGAGAGCGTCTCCAGCATGACCAAGCAAATCTGCTTGACACACTCGGTGACAGACTGCGGCACG
 20 CCAGCGATGGTGTATGNGCCCGCTCGGTGGAGTTGGGCAGCATATCCCCCGGCACCTGGANCTGCGCCCCCGTACT
 CTCGCGATCTCTTTGATCTTACACCCGCTTTCCCAATCAGTGGAGCCGCACTGGGTGTGCTGGCACCACCAG
 CCTCAGGGTGACCGGGGTGCTGCTGGCCGCGTACTGTTGGTCATGGAGCTGTTGATATCTTCTCCAGCTTG
 TCGATGATCATAGCGAAAGCCTTAAAGATGGCATTTGGTGGGGCCGGTCAGAGTGATGATTCTCTCCGACAATTC
 CCTCCGAGATGTTGATCCGCGCGCAATCTCCTCGCGGATCCTCTTAACGAATCCCCTTTCTTCCAAANGATGAT
 25 TCCTANTTCTTTACCGTGCATATAGAAAGCCCGAAATGGTTGACGAGTCCACATTTAGTCCACACTTTTCAGTCA
 CACCCGGAATCCAATGGGCGAGCGGAGGGACGGAGACGGGGGAAGATGGGCGCCGAACGTTGGGCCAAGAATA
 GAGCAAGAGAGNAGAAGGAAAAANGACAAAGAGCACAAACAGAAAAAAGACAAAAAGGNAGCAAGCACACNAAAA
 AA

SEQ ID NO:31

GL2_209_2_M13F

CGCTGCGGCCCTCCCGCCCGCGCTCGCCATGGATGCCGGTGTGACTGAAAGTGGACTAAATGTGACTCTCACCAT
 TCGGCTTCTTATGCACGGAAGGAAGTAGGAAGCATCATTTGGGAAGAAAGGGGAGTCGGTTAAGAGGATCCGCGA
 GGAGAGTGGCGCGCGGATCAACATCTCGGAGGGGAATTGTCCGGAGAGAATCATCACTCTGACCGGCCCCACCAA
 35 TGCCATCTTTAAGGCTTTTCGCTATGATCATCGACAAGCTGGAGGAAGATATCAACAGCTCCATGACCAACAGTAC
 CGCGGCCAGCAGGCCCCCGGTCACCTGAGGCTGGTGGTGCCGGACACCCAGTGGGCTCCCTGATTGGGAAAGG
 CGGGATGTAAGATCAAAG

SEQ ID NO:32

>gi|14141164|ref|NM_006196.2| Homo sapiens poly(rC) binding protein 1
 (PCBP1), mRNA

GACGCCGCCCGACCTGCGACTACGCTGCGGACTCCCGCCCGCTCCCGCTCGCTCCCGCGGTCTCTCGCTC
 GCCTCGCCCGGTAGTTTTGGGCCTACACCTCCCCCTCCCCCGCCAGCCGCAAAGACTTGACCACGTAA
 CGAGCCCCAACTCCCCCGAACGCCGCCCGCGCTCGCCATGGATGCGGGTGTGACTGAAAGTGGACTAAAT
 45 GTGACTCTCACCATTTCGGCTTCTTATGCACGGAAGGAAGTAGGAAGCATCATTTGGGAAGAAAGGGGAGT
 CGGTAAAGAGGATCCGCGAGGAGAGTGGCGCGCGGATCAACATCTCGGAGGGGAATTGTCCGGAGAGAAT
 CATCACTCTGACCGGCCCCACCAATGCCATCTTTAAGGCTTTTCGCTATGATCATCGACAAGCTGGAGGAA
 GATATCAACAGCTCCATGACCAACAGTACCGCGGCCAGCAGGCCCCCGGTACCCCTGAGGCTGGTGGTGC
 CGGCCACCCAGTGCGGCTCCCTGATTGGGAAAGGCGGGTGTAAAGATCAAAGAGATCCGCGAGAGTACGGG
 50 GGCGCAGGTCCAGGTGGCGGGGGATATGCTGCCCAACTCCACCGAGCGGGCCATCACCATCGCTGGCGTG
 CCGCAGTCTGTACCGAGTGTGTCAAGCAGATTTGCTGGTTCATGCTGGAGACGCTCTCCAGTCTCCGC
 AAGGAGAGTCATGACCATTCCTGACCGCCATGCGCGCCAGCTCCCGAGTCATCTGCGCGGGCGGCCA
 AGATCGGTGCAGCGACGCTGTGGGCTACCCCCATGCCACCCATGACCTGGAGGGACCACCTCTAGATGCC
 TACTCGATTCAAGGACAACACACCATTTCTCCGCTCGATCTGGCCAAGCTGAACCAGGTGGCAAGACAAC
 55 AGTCTCACTTTGCCATGATGCACGGCGGGACCGGATTGCGCGGAATTGACTCCAGCTCTCCAGAGGTGAA
 AGGCTATTGGGCAAGTTTGGATGCATCTACTCAAACCCCATGAACTCACCATTCCAAATAACTTAATT
 GGCTGCATAATCGGGCGCCAAGGCGCCAACATTAATGAGATCCGCCAGATGTCCGGGGCCAGATCAAAA
 TTGCCAACCCAGTGGAAGGCTCCTCTGGTAGGCAGGTTACTATCACTGGCTCTGCTGCCAGTATTAGTCT
 GGCCCAGTATCTAATCAATGCCAGGCTTTCTCTGAGAAGGGCATGGGGTGCAGCTAGAACAGTGTAGGT
 60 TCCCTCAATAACCCCTTTCTGCTGTTCTCCCATGATCCAATGTGTAATTTCTGGTCAGTGATTCCAGGT

TTTAAATAATTTGTAAGTGTTCAGTTTCTACACAACCTTTATCATCCGCTAAGAATTTAAAAATCACATTC
TCTGTTTCAGCTGTTAATGCTGGGATCCATATTTAGTTTTATAAGCTTTTCCCTGTTTTTAGTTTTGTTTT
GGGTTTTTTGGCTCATGAATTTTATTTCTGTTTGTGCGATAAGAAATGTAAGAGTGGAATGTTAATAAATT
TCAGTTTAGTTCTGTAATGTCAAGAATTTAAGAATTAACAAACGATTGGTTAAAAAATGCTTCATATTT
5 GAAAAAGCTGGGAATTGCTGTCTT

SEQ ID NO:33

>gi|5453854|ref|NP_006187.1| poly(rC)-binding protein 1; heterogenous
nuclear ribonucleoprotein X; alpha-CP1; nucleic acid binding protein sub
10 2.3; heterogenous nuclear ribonucleoprotein E1 [Homo sapiens]
MDAGVTESGLNVTLTIRLLMHGKEVGSIIIGKKGESVKRIRESGARINISEGNCPERIITLTGPTNAIFK
AFAMIIDKLEEDINSSMTNSTAASRPPVTLRLVVPATQCSLIGKGGCKIKEIRESTGAQVQVAGDMLPN
STERAITIAGVPQSVTECVKQICLVMLETLQSPPQGRVMTIPYQMPASSPVICAGGQDRCSDAVGYPHA
THDLEGPPLDAYSIQGQHTISPLDLAKLNQVARQQSHFAMMHGGTGFAGIDSSPEVKGYWASLDASTQT
15 THELTIPNNLIGCIIGROGANINEIRQMSGAQIKIANPVEGSSGRQVTITGSAASISLAQYLINARLSSE
KGMGCS

SEQ ID NO:34

Amino acid transporter (SLC1A5)

>G3_8_66_PCR_G3F1
CTCGCCCGGTGAAGACGAAGTGCCTCAAGCGCTCCAGTGCCCAACGCCAGCGCACCCCCGGCCCCCGACACCCC
CAGTCCCACGCCACGCCCGCCACCACGGCACCCTTACAACAACGGTTGCTCAAGCAGCGGCACCTTGTCCTGGGA
ACCGCAGTAGCCGCTGCTGCCGCGCTTGGTCTCGATGGAGGCCAGCGCCAGGCCCCGTTGGCGGTGGCTCC
25 GCCGCTCGAGCCCCCTTGGAGTCTCGAGGAGGATCGGCCACCATGATGGAAGCACCGGGGTTTCTTAGCGCTGGA
AGCTGGCTGGGAGCGCTTGGGCTCCTTCCCAGGACCCGACGTTCTTAGGACTGAGTTGAGTAACAGCACCTGGAG
ACTGGAACCTTTGGAGGGCTCCTTAGAGTTGTGAGTTCACAGCACTAAGTTCCTTGGCTCTTGAAGCTGGAGTGT
TTAAATTCCTCAGGCTGGGCGCTAGGCTTCTC

SEQ ID NO:35

>GL3_8.66_A_M13F

CTCGCCCGGGAAGACGAAGGCGCTCAAGCGCTCCAGGCCCAACGCCAGCGCACCCCCGGCCCCCGACACCCCCAG
TCCAGCGCCACGCCCGCCACCACGGCCACCACTGTGAGCAGCACAAAGCAGTTGGCTCGAAGGCAGCGGCGCAC
CTGGTCCCGGGAACCCGAGTCCGCGCCCTGCTGCCGCGCTTGGTCTCGATGGAGGCCAGCGCCAGGCCCCGTT
35 GGCGGTGGGCTCCGCGCTGCGAGCCCCCTTGGAGTCTCGAGGAGGATCGGCCACCATGATGGGAAGCACCGGGGT
TTCTTAGCGCTGGAAGCTGGCTGGGAGCGCTTGGGCTCCTTCCCAGGACCCGACGTTCTTAGGACTGAGTTGAG
TAACAGCACCTGGAGACTGGAACCTTTGGAGGGCTCCTTAGAGTTGTGAGTTCACAGCACTGAAGTTCCTTGGCTC
TTGGAAGCTGGAGTGTTTAAATTCCTCAGGCTGGGCGCTGAGGCTTCTCTGCTCTGCCCCGTGTGCCAGATGTCC
GAAAGCTGGGAGTTCGAGCGCCCCGGGTTCCT

SEQ ID NO:36

>GL3_8.66_A_M13R Direction: anti-sense

CTCCTAGGGCCAAGGAACCCGGGCGCTCCGAACTCCAGCTTTCGGACATCTGGCACACGGGGCAGAGCAGAGAA
GCCTCAGCGCCAGCCTGGGGAATTTAAACACTCCAGCTTCCAAGAGCCAAGGAACCTCAGTGCTGTGAATCAC
45 AACTCTAAGGAGCCCTCCAAAGTTCCAGTCTCCAGGTGCTGTTACTCAACTCAGTCTTAGGAACGTCGGGTCCTG
GGAAGGAGCCCAAGCGCTCCAGCCAGCTTCCAGGCGCTAAGAAACCCCGGTGCTTCCCATCATGGTGGCCGATC
CTCCTCGAGACTCCAAGGGGCTCGCAGCGGCGGAGCCACCGCCAACGGGGGCTGGCGCTGGCCTCCATCGAGG
ACCAAGGCGCGGCAGCAGGCGGCTACTGCGGTTCCTGGGACCAGGTGCGCCGCTGCCTTCGAGCCAACCTGCTTG
TGCTGCTGACAGTGGTGGCCGTGGTGGCCGCGTGGCGCTGGGACTGGGGGTGTGCGGGGCGGGGGTGGCGCTGG
50 CGTGGGCTGGAGCGCTTGAGCGCTTCGTCTTCCCGGGCGAGCCGAACCCAGCACACTGGAGCCACC

SEQ ID NO:37

>GL3_8.66_B_M13F Direction: sense

CTATAGGGCGAATGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCT
TCTCTCGGCATGAGCAGCTGTACAAGGAGGAGGCCCAAGGCCGTTGGCAGCGGTGGCTCCAGTTGTGCTGGGT
55 TCCGCTCGCCCGGGAAGACGAAGGCGCTCAAGCGCTCCAGGCCCAACGCCAGCGCACCCCCGTTCCCGACACC
CCCAGTCCCAGCGCCACGCCGCGCCACCACGGCCACCCTGCTCAGCAGCACAAGCAGGTTGGCTCGAAGGCAGCGG
CGCACCTGGTCCCGGGAACCGCAGTAGCCGCTGCTGCCGCGCTTGGTCTCGATGGAGGCCAGCGCCAGGCCC
CCGTTGGCGGTGGGCTCCGCCGCTGCGAGCCCCCTTGGAGTCTCGAGGAGGATCGGTACCATGATGGGAAGCACC
GGGTTTCTTAGCGCTGGAAGCTGGCTGGGAGCGCTTGGGCTCCTTCCCAGGACCCGACGTTCTTAGGACTGAG
60 TTGAGTAACAGCACCTGGAGACTGGAACCTTTGGAGGGCTCCTTAGAGTTGCGAGTTCACAGCACTGAAGTTCCTT

GGCTCTTGGAAGCTGGAGTGTAAATTTCCCAAGGCTGGGCGCTGAGGCTTCTCTGCTCTGCCCCGTGTGCCAGA
TGTCCGAAAGCTGGGAGTTCGGAGCGCCCGGGTTCCTTGGCCCTAGGAG

SEQ ID NO:38

5 >GL3_8.66_B_M13R Direction: anti-sense
CTCCTAGGGCCAAGGAACCCGGGCGCTCCGAAGTCCAGCTTTCGGACATCTGGCACACGGGGCAGAGCAGAGAA
GCCTCAGCGCCAGCCTGGGGAATTTAAACACTCCAGCTTCCAAGAGCCAAGGAACCTTCAGTGCTGTGAACTCGC
AACTCTAAGGAGCCCTCCAAAGTTCCAGTCTCCAGGTGCTGTTACTCAACTCAGTCCTAGGAACGTCGGGTCTCTG
10 GGAAGGAGCCCAAGCGTCCAGCCAGCTTCCAGGCGCTAAGAAACCCCGGTGCTTCCCATCATGGTGACCGATC
CTCCTCGAGACTCCAAGGGGCTCGCAGCGGCGGAGCCACCGCCAACGGGGGCTGGCGCTGGCCTCCATCGAGG
ACCAAGGCGCGCAGCAGGCGGCTACTGCGGTTCCCGGGACCAGGTGCGCCGCTGCCTTCGAGCCAACCTGCTTG
TGCTGCTGACAGTGGTGGCCGTGGTGGCCGGCGTGGCGCTGGGACTGGGGGTGTGCGGGACCGGGGTGCGCTGG
CGTTGGGCTTGAGCGCTTGAGCGCCTTCGTCTTCCCGGGCGAGCCGGAACCCAGCACACTGGAGCCACCGCTGC
15 CACCGGCTTGGCGGCTCCTCCTTGACAGCTCGTCCATGCCGAGAGAAGGGCGAATTCTGCAGATATCCATCA
CACTGG

SEQ ID NO:39

>GL3_8.66_C_M13F Direction: anti-sense
CTCCTAGGGCCAAGGAACCCGGGCGCTCCGAAGTCCAGCTTTCGGACATCTGGCACACGGGGCAGAGCAGAGAA
20 GCCTCAGCGCCAGCCTGGGGAATTTAAACACTCCAGCTTCCAAGAGCCAAGGAACCTTCAGTGCTGTGAACTCAC
AACTCTAAGGAGCCCTCCAAAGTTCCAGTCTCCAGGTGCTGTTACTCAACTCAGTCCTAGGAACGTCGGGTCTCTG
GGAAGGAGCCCAAGCGTCCAGCCAGCTTCCAGGCGCTAAGAAACCCCGGTGCTTCCCATCATGGTGGCCGATC
CTCCTCGAGACTCCAAGGGGCTCGCAGCGGCGGAGCCACCGCCAACGGGGGCTGGCGCTGGCCTCCATCGAGG
ACCAAGGCGCGCAGCAGGCGGCTACTGCGGTTCCCGGGACCAGGTGCGCCGCTGCCTTCGAGCCAACCTGCTTG
25 TGCTGCTGACAGTGGTGGCCGTGGTGGCCGGCGTGGCGCTGGGACTGGGGGTGTGCGGGGCGCGGGGTGCGCTGG
CGTTGGGCTTGAGCGCTTGAGCGCCTTCGTCTTCCCGGGCGAGCCGGAACCCAGCACACTGGAGCCACCGCTGC
CACCGGCTTGGCG

SEQ ID NO:40

30 >GL3_8.66_C_M13R Direction: sense
GATCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTCTCTCGGCATGGACGAGCTGTACAAGGAGGAGGC
CGCCAAGGCCGGTGGCAGCGGTGGCTCCAGTGTGCTGGGTTCCGGCTCGCCCCGGGAAGACGAAGGCGCTCAAGCG
CTCCAGGCCCAACGCCAGCGCACCCCCGGCCCCCGACACCCCCAGTCCAGCGCCACGCCGGCCACCACGGCCAC
CACTGTCTCAGCAGCACAAAGCAGGTTGGCTCGAAGGCAGCGCGCACCTGGTCCCGGGAACCGCAGTAGCCGCCTGC
35 TGCCGCGCCTTGGTCTCGATGGAGGCCAGCGCCAGGCCCCCGTTGGCGGTGGGCTCCGCGCTGCGAGCCCCCTT
GGAGTCTCGAGGAGGATCGGCCACCATGATGGGAAGCACCGGGGTTTCTTAGCGCCTGGAAGCTGGCTGGGAGCG
CTTGGGCTCCTTCCAGGACCCGACGTTCTTAGGACTGAGTTGAGTAACAGCACCTGGAGACTGGAACCTTTGGAG
GGCTCCTTAGAGTTGTGAGTTCACAGCACTGAAGTTCTTGGCTCTTGGAAAGCTGGAGTGTAAATTTCCCAAGG
CTGGGCGCTGAGGCTTCTCTGCTCTGCCCCGTGTGCCAGATGTCCGAAAGCTGGGAGTTCGGAGCGCCCGGGTTC
40 CTTGGCCCTAGGAG

SEQ ID NO:41

>GL3_8.66_D_M13F Direction: N/A
ACTATAGGGCGAATTGGGCCCTCTAGATGCATGCTCGAGCGGCCCGCCAGTGTGATGGATATCTGCAGAATTCGCC
45 CTTCTCTCGGCATGGACGAGCTGTACAAGGAGGAGGCCCGCAAGGCCGGTGGCAGCGGTGGCTCCAGTGTGCTGG
GTTCCGGCTCGCCGGAAGACGAAGGCGCTCAAGCGCTCCAGGCCCAACGCCAGCGCACCCCCGGCCCCCGACA
CCCCAGTCCAGCGCCACGCCGGCCACCACGGCCACCACTGTCTCAGCAGCACAAAGCAGGTTGGCTCGAAGGCAGC
GGCGCACCTGGTCCCGGGAACCGCAGTAGCCGCTGCTGCCGCGCTTGGTCTCGATGGAGGCCAGCGCCAGGC
CCCCGTTGGCGGTGGGCTCCGCGCTGCGAGCCCCCTTGGAGTCTCGAGGAGGATCGGCCACCATGATGGGAAGCA
50 CCGGGGTTTCTTAGCGCCTGGAAGCTGGCTGGGAGCGCTTGGGCTCCTTCCAGGACCCGACGTTCTTAGGACTG
AGTTGAGTAACAGCACCTGGAGACTGGAACCTTGGAGGGCTCCTTAGAGTTGTGAGTTCACAGCACTGAAGTTC
TTGGCTCTTGGAAAGCTGGAGTGTAAATTTCCCAAGGCTGGGCGCTGAGGCTTCTCTGCTCTGCCCCGTGTGCCA
GATGTCCGAAAGCTGGGAGTTCGGAGCGCCCGGGTTCCTTGGCCCTAGGAGCCGC

55 SEQ ID NO:42

>GL3_8.66_D_M13R Direction: anti-sense
CTCCTAGGGCCAAGGAACCCGGGCGCTCCGAAGTCCAGCTTTCGGACATCTGGCACACGGGGCAGAGCAGAGAA
GCCTCAGCGCCAGCCTGGGGAATTTAAACACTCCAGCTTCCAAGAGCCAAGGAACCTTCAGTGCTGTGAACTCAC
AACTCTAAGGAGCCCTCCAAAGTTCCAGTCTCCAGGTGCTGTTACTCAACTCAGTCCTAGGAACGTCGGGTCTCTG
60 GGAAGGAGCCCAAGCGTCCAGCCAGCTTCCAGGCGCTAAGAAACCCCGGTGCTTCCCATCATGGTGGCCGATC

CTCCTCGAGACTCCAAGGGGCTCGCAGCGGCGGAGCCCCACCGCCAACGGGGGCTGGCGCTGGCCTCCATCGAGG
ACCAAGGCGCGGCAGCAGGCGGCTACTGCGGTTCCCGGGACCAGGTGCGCCGCTGCCCTTCGAGCCAACCTGCTTG
TGCTGCTGACAGTGGTGGCCGTGGTGGCCGGCGTGGCGCTGGGACTGGGGGTGTCGGGGGCGGGGGTGCCTGG
CGTTGGGCTTGGAGCGCTTGGAGCGCTTCGTCTTCCGGGCGAGCCGGAACCCAGCACACTGGAGCCACCGCTGC
5 CACCGGCCTTGGCGGCCTCCTCCTTGTACAGCTCGTCCATGCCGAGAGAAGGGCGAATTCTGCAGATATNCATCA
CACTGGCG

SEQ ID NO:43

>gi|5032092|ref|NM_005628.1| Homo sapiens solute carrier family 1 (neutral
10 amino acid transporter), member 5 (SLC1A5), mRNA
GTAACCGCTACTCCCGGACACCAGACCACCGCCTTCCGTACACAGGGGCCCCGCATCCCACCCTCCCGGAC
CTAAGAGCCTGGGTCCCCGTGTTTCCGGAGGTCCGCTTCCCGGCCCCAGATTCTGGCATCCCAGCCCTCA
GTGTCCAAAGACCCAGGCAGCCCGGGTCCCCGCCTCCCGGATCCAGGCGTCCGGGATCTGCGCCACCAGAA
15 CCTAGCCTCCTGCAGACCTCCGCCATCTGGGGGCACTCAACCTCCTGGAGCCAAGGGGCCCCACGTCCAC
CCAGAGAAACTCTCGTATTCAGCTCCTAGGGCCAAGGAACCCGGGCGTCCGAACCTCCAGCTTTCGG
ACATCTGGCACACGGGGCAGAGCAGAGAAGCTCAGCGCCAGCCTGGGGAATTTAAACACTCCAGCTTCC
AAGAGCCAAGGAACCTCAGTGTCTGTAACCTCAAACTTAAGGAGCCCTCCAAAGTTCAGTCTCCAGGT
GCTGTTACTCAACTCAGTCTTAGGAACGTGGGTCTGGGAAGGAGCCCAAGCGCTCCAGCCAGCTTCC
20 AGGCGCTAAGAAACCCGGTGTCTCCCATCATGGTGGCCGATCCTCCTCGAGACTCCAAGGGGCTCGCAG
CGGCGGAGCCCCACCGCCAACGGGGGCTGGCGCTGGCCTCCATCGAGGACCAAGGCGCGGACAGGCGG
CCTATGCGGTTCCCGGGACAGGTGCGCCGCTGCCTTCGAGCCAACCTGCTTGTGCTGCTGACAGTGTG
GCGTGGTGGCCGGCTGGCGCTGGGACTGGGGGTGTCGGGGGCGGGGGTGCCTGGCGTGGGCCCCGG
AGCGCTTGAGCGCTTCGTCTTCCCGGCGAGCTGCTGCTGCGTCTGCTGCGGATGATCATCTTGCCGCT
GGTGGTGTGACGCTTGATCGGCGGCGCCGACGCTGGACCCGGGCGCTCGGCCGTCTGGGCGCCTGG
25 GCGCTGCTCTTTTCTGGTCAACACGCTGCTGGCGTGGCGCTCGGAGTGGGCTTGGCGCTGGCTCTGC
AGCCGGGCGCCGCTCCGCCGCCATCAACGCCTCCGTGGGAGCCGCGGGCAGTGCCGAAAATGCCCCAG
CAAGGAGGTGCTCGATTGCTTCCCTGGATCTTGCGAGAAATATCTTCCCTTCCAACCTGGTGTGACAGCC
TTTCGCTCATACTCTACCACCTATGAAGAGAGGAATATCACCGGAACAGGGTGAAGGTGCCCGTGGGGC
AGGAGGTGGAGGGGATGAACATCCTGGGCTTGGTAGTGTGTTGCCATCGTCTTTGGTGTGGCGCTGCGGAA
30 GTGGGGCTGAAGGGGAGCTGCTTATCCGCTTCTTCACTCCTTCAATGAGGCCACCATGGTCTTGCTG
TCCTGGATCATGTGTTACGCCCTGTGGGATCATGTTCTGCTGGTGGTGGCAAGATCGTGAGATGGAGG
ATGTGGGTTTACTCTTTGCCCGCTTGGCAAGTACATTCTGTGCTGCTGCTGGGTACGCCATCCATGG
GCTCCTGGTACTGCCCCCTCATCTACTTCTTCAACCGCAAAAACCCCTACCGCTTCTGTGGGGCATC
GTGACGCGCTGGCCACTGCTTTGGGACCTCTTCCAGTTCGCCACGCTGCCGCTGATGATGAAGTGG
35 TGGAGGAGAATAATGGCGTGGCCAAGCACATCAGCCGTTTCATCCTGCCCATCGGCGCCACCGTCAACAT
GGACGGTGCCGCGCTCTCCAGTGGCGTGGCCGAGTGTTTATTGCACAGCTCAGCCAGCAGTCTTGGAC
TTCGTAAAGATCATCACCATCCTGGTACGGCCACAGCGTCCAGCGTGGGGGCGAGCGGGCATCCCTGCTG
GAGGTGCTCTACTCTGGCCATCATCCTCGAAGCAGTCAACCTCCCGGTGACCATATCTCCTTGATCCT
GGCTGTGACTGGCTAGTTCGACCGGCTGTACCGTCTCAATGTGTAAGGTGACGCTCTGGGGGCGAGGA
40 CTCCTCCAAAATTTATGTGGACCGTACGGAGTGCAGAAGCACAGAGCCTGAGTTGATACAAAGTGAAGAGTG
AGCTGCCCCCTGGATCCGCTGCCAGTCCCCACTGAGGAAGGAACCCCTCCTCAAACTATCGGGGGCC
CGCAGGGGATGCCACGGTCGCCCTCTGAGAAGGAATCAGTCATGTAAACCCGGGAGGGACCTTCCCTGCC
CTGCTGGGGGTGCTCTTTGGACACTGGATTATGAGGAATGGATAAATGGATGAGCTAGGGCTCTGGGGGT
CTGCCTGCACACTCTGGGGAGCCAGGGGCCCCAGCACCTCCAGGACAGGAGATCTGGGATGCCTGGCTG
45 CTGGAGTACATGTGTTTCAAGGGTTACTCCTCAAAACCCCAAGTCTCACTCATGTCCCCAACCTCAAGG
CTAGAAAACAGCAAGATGGAGAAATAATGTTCTGCTGCGTCCCCACCGTGACCTGCCTGGCCTCCCTGT
CTCAGGGAGCAGGTACAGGTACCATGGGGAATTCTAGCCCCACTGGGGGGATGTTACAACACCATGC
TGGTTATTTTGGCGGTGTAGTTGTGGGGGGATGTGTGTGTGTCACGTTGTGTGTGTGTGTGTGTGTGTGT
TGTTGTGTGTGTCTGTGACCTCCTGTCCCCATGGTACGTCCCACCGTGTCCCAGATCCCCATTCCCTC
50 CACAATAACAGAAACACTCCCAGGGACTCTGGGGAGAGGCTGAGGACAAATACCTGCTGTCACTCCAGAG
GACATTTTTTTAGCAATAAAATTGAGTGTCAACTATTTAAAAA

SEQ ID NO:44

>gi|5032093|ref|NP_005619.1| solute carrier family 1 (neutral amino acid
55 transporter), member 5; baboon M7 virus receptor; RD114 virus receptor;
neutral amino acid transporter B [Homo sapiens]
MVADPPRDSKGLAAAEPTANGGLALASIEDQAAAGGYCGSRDQVRRCLRANLLVLLTVVAVVAGVALGL
GVSGAGGALALGPERLSAFVFPGELLRLRLMIILPLVVCSLIGGAASLDPGALGRLGAWALLFLVTTL
LASALGVGLALALGPGAASAINASVGAAGSAENAPSKEVLDSFLDLARNIFPSNLVSAAFRSYSTTYEE
60 RNITGTRVKVPVQGVEGMNIGLVVFAIVFGVALRKLGPGEGLLIRFFNSFNEATMVLVSWIMWYAPVG
IMFLVAGKIVEMEDVGLLFARLGKYLCLLGHAIHGLLVPLIYFLFTRKNPYRFLWGIIVTPLATAFGT

SSSSATLPLMMKCVEENNGVAKHISRFLPIGATVNMDGAALFQCVAAVFIAQLSQSLDFVKIITILVT
 ATASSVGAAGIPAGGVLTLAILLEAVNLPVDHISLILAVDWLVDRSCTVLNVEGDALGAGLLQNYVDRTE
 SRSTEPELIQVKSELPLDPLVPTEEGNPLLKHRYRGPAGDATVASEKESVM

5 SEQ ID NO:45

Chromobox homologue 6

>GL3_37C_2_M13F

ACGCAGGACGCTCTCGCTGAACCTTCTTGCTCTTGCCCTATAACGCGGTTCCGCGAGGGGACTTTGGGGCGGGCCAG
 CGCCCCGGCCCCCTGCCCCGGGCCCCCGCCGAGCGCCCTTTGTGCGATCACCTTCAGGTCAGGATGATGCGGTCCG
 10 CTTGGGCATCCCGCNCCTTACACCTTGCGGTTGATGATGCGCACCGTCTCCGAGAAGGGCGAAATGAGGCGGGCA
 GCAGTCCGGGGCTGCCCCCTGCAGGTTCCGGGCGGGGGCAGGGGACGGCGGGACATACGGTTGGGCAAGNCGGG
 CCGGCGAATGTTCTTATACCTTGAAGCCGGGTTGCCACGGGATTAGACGCCTTTGGGAGATTGCACGCATATGG
 GGCAAGGAAGGCGNACCTTGCGCCGCTCGGGCNTTNGACACGGAAGGNAATTTGCCACATTCCACTTGAATT
 15 TGCGCCAGAATANGGCGNCCCACTTNCGTGAGCGACCACTTTGGGCCCCCACGGGGANGGNACGCCGTTTATCA
 TCAAAGAGGGGAAAAGGCGNAAAAAAGCGTTNTATCAGTAGAGGGGACAGGGGAATTATTAGGGGGGATTAC
 CCCCCCAGTTACAATTTTACCATATTGAGGAGGAGNCCGCCCANAGTATTAAACATAGGACGTGTTNCAAACG
 GGGACNTTACACCAGCATTANCAATTTCCGCAATTTATTATAGTATGGGGATTTCNAGAACAAANGTGGGCGCTTT
 TGTGCAAAAATTTAGGAATAGGTCCCCACGG

20 SEQ ID NO:46

>GL3_35L_1_M13R Direction: anti-sense

ACAGCACTTGGGAGCCCGAGGAGAATCCTTGACTCGCGGCTCATTGCAGCCTTCGAACAAAAGGAGAGGGAGC
 GTGAGCTGTATGGGCCCAAGAAGAGGGGACCCAAACCCAAACTTTCTCCTGAAGGCGCGGGCCAGGCCGAGG
 CCTNTCCGCATCAGTGTATGTGCAATTTCTCTGTCAAGCCGAGCGCCAGTGCCTCCTCGCCCAAGCTGCACTCCAG
 25 CGCAGCCGTGCACCGGCTCAAGAAGGACATCCGCCGCTGCCACCGTATGTCCCGCCGTCCCCTGCCCCGCCCCGA
 CCCGAGGGGGGCGAGCCCCGACTGCGCCCGCCCATTTGCGCCTTCTCGGAGACGGTGCATCATCAACCGCCA
 AGGTGAAGCCGCGGGAGCCCAAGGCGGAACCGCATCATCTGAACCTGAAGGTGATCGGAACAAGGGCGCTGGCA
 CGGCAGGGGGCGCCGGGCAAGGGGGCGGGGACGCTGGCCCGCCAAAAAAGGNTCNCCTATNGCGGGGAAAC
 CCGGCGGTGTAATAGGGGGCAAGAGGCAAGAACAGGTACCAGGCAAGGAAGCCGGGTACATGGCAGGTNCGCG
 30 GNCANAGAAACCCAGCAACACTTGGGAGGCCACACCGGTNGCNCAACCGGACCTTTNGGGGGGGCCTCACATC
 ACCNATGGTACCAAGGGTTCGTACCAATCGAGCACGGACAAGAAAGGGGGCGAAATTC

SEQ ID NO:47

>GL3_35L_3_M13R Direction: anti-sense

ACAGCACTTGGGAGCCCGAGGAGAATCCTTGACTCGCGGCTCATTGCAGCCTTCGAACAAAAGGAGAGGGAGC
 GTGAGCTGTATGGGCCCAAGAAGAGGGGACCCAAACCCAAACTTTCTCCTGAAGGCGTGGGCCAGGCCGAGG
 CCTTCCGCATCAGTGTATGTGCAATTTCTCTGTCAAGCCGAGCGCCAGTGCCTCCTCGCCCAAGCTGCACTCCAGCG
 CAGCCGTGCACCGGCTCAAGAAGGACATCCGCCGCTGCCACCGTATGTCCCGCCGTCCCCTGCCCCGCCCCGACC
 CGCAGGGGGGCAAGCCCCGGAACCTGNGCCCGCCCAATTTGCGCCTTCTCGGAGACCGGTGCGCATCATCAACCGC
 40 AAGTGAAGCCGCGGGAGCCAAGCGGAACCGATCATCTGAAACCTGAAGGTGATCGGACAAGGGCGCTGGCGGAC
 GGGGGCGCCGGGCGAGGGGGGCGGGGGCGCCTGGGCCCCGGCCCCCAAAANAGTCCCCCACCATCAGGGCGAGG
 GAAAACCCCGGGAGGANAAAGTAGCGGGCAAAANGAAGCCAANAGGAAANGTTCAACAAGGAGAGAAGAACNGA
 CGACCANGGCGAGATCCCGGGGAAGAGAAACGCNNAAGGCAAAACAAGATTGGGAGAAGCNACACNNGGTGAGC
 ACAACAGCCGGGGCCACTAGGCGGCGGNACCTACCCTCCCCTATTGGAACGCAGCNCNCGGGCANCAATTGCCGA
 45 GAAAGAAAGGGGGGCGAATCTNCGA

SEQ ID NO:48

>GL3_35L_4_M13F Direction: N/A

ACGCAGGACGCTCTCGCTGAACCTTCTTGCTCTTGCCCTATAACGCGGTTCCGCGAGGGGACTTTGGGGCGGGCCAG
 50 CGCCCCGGCCCCCTTGCCCGGCGCCCCCGCCGCGCCAGCGCCCTTGTGCGATCACCGTTTCAGGTTTCAGGATGATGCG
 GTTCCGGCTTTGGGCATCCCGGGGGGCTTACCTTTGCNNGGTTGACTGATGCGCACCGGACATCTGAGAAAGGG
 CGAAATGGGACGGGACGCAGATCCGGGGCNTGCCCCCCTTGCCGGGGGGACTTCCAGGGGCACGGGGGANGA
 CAAGGGGGGAACGGGGACGGGTGAACCAATTACCGGNGGTGGGCCAGCCGGGCGGGGAATGGGTCCCTTCTTGGA
 AGACCGGGTGCCAA

55

SEQ ID NO:49

>GL3_37C_1_M13F Direction: sense

ACGCAGGACGCTCTCGCTGAACCTTCTTGCTCTTGCTTATAACGCGGTTCCGCGAGGGGACTTTGGGGCGGGCCAG
 60 CGCCCCGGCCCCCTTGCCCTGGAGCCCCCGCCGCGCCAGCAGGCCCATTTGTGAGAATTCAACCTTTCAAGGNTNCAAG
 GGATGAATTGCGNGGTCCCGGCTNTATGGGGATTCCNCAAGAGGGACTTTTCAACCTTTTGCCAGGGATATGGAA

CGTGGNATTTGCCGGCCAAACCCCGGGGTNACTCCCCAGAAAGGAAAAAGGGGGCGGGAAAAAATTTGGGCGCT
GGGGGCCGACGGC

SEQ ID NO:50

5 >GL3_35L_1_M13F Direction: N/A

ACGCAGGACGCTCTCGCTGAACTTCTTGCTCTTGCCCTATAACGCGGTTCCGCGAGGGGACTTTGGGGCGGGCCA
GCGCCCGGGCCCCCTTGCCCCGGGGCCCCCGCCCGCCNAGCGGCCCTTTGTTCGGATCAACCTTCAGGTCAGGATGA
TGCGGTTCCGCTTTGGGCTTCCCGCGGGTTTACCTTGCGGGTTGATGATGCGGCAACCCGGGTCTTCCCAGA
AGAACAGGGGGCGGACAATTGGNGCGCAGGGGC

SEQ ID NO:51

10 >GL3_35L_3_M13F Direction: N/A

ACGCAGGACGCTCTCGCTGAACTTCTTGCTCTTGCCCTATAACGCGGTTCCGCGAGGGGACTTTGGGGCGGGCCAG
CGCCCCGGCCCCCTTGCCCCGGGGCCCCCGCCCGCCAGCGCCCTTGTCGATCACCTTCAGGTNCAGGATGATGCGGTT
15 CCGCTTGGGCTCCCGCGGCTTACCTTGCGGTTGATGATGCGCACCGTCTCCGAGAAGGGCGAAATGGGCGGGCG
CAGTCCGGGGCTGCCCCCTTGCGGGATCCGGGCGGGGCAGGGGACCGGCGGGCACCATACGGTGGCAGCGGGC
GGAATGTCTTCTGAAGCCGGTGACGGCTTGCGCTTGAAGTTGCAGCTTGGGCCAAGG

SEQ ID NO:52

20 >GL3_35L_4_M13R Direction: anti-sense

ACAGCACTTGGGAGCCCCGAGGAGAACATCCTGGACTCGCGGCTCATTGCAGCCTTCGAACAAAAGGAGAGGGAGC
GTGAGCTGTATGGGCCCAAGAAGAGGGGACCCAAACCCAAAACCTTCTCTCTGAAGGCGCGGGCCCAGGCCGAGG
CCCTCCGCATCAGTGATGTGCATTTCTCTGTCAAGCCGAGCGCCAGTGCTCTCTCGCCCAAGCTGCACTCCAGCG
CAGCCGTGCACCGGCTCAAGAAGGACATCCGCCGCTGCCACCGTATGTCCCGCGTCCCTTGCCCCCGCCGGACC
25 CGCAGGGGGGCGAGCCCCGACTGCGCCCGCCATTTCGCCCTTCTCAGAGACGGTGCGCATCATC

SEQ ID NO:53

30 >GL3_35L_PCR_G3F1 Direction: N/A

GACGCAGGACGCTCTCGCTGAACTTCTTGCTCTTGCCCTATAACGCGGTTCCGCGAGGGGACTTTGGGGCGGGCC
30 AGCGCCCCGGCCCCCTGCCCCGGGGCCCCCGCGCGGCCCTTTATACCTTCAGTTCAGATGATGCGGTTCCGCTTGGG
CTCCCGCGGCTTACCTTGCGGTTGATGATGCGCACCGTCTCCGAGAAGGGCGAAATGGGCGGGCGCAGTCCGGG
GCTGCCCCCTGCGGGTCCGGGCGGGGCAGGGGACGGCGGGACATACGGTGGCAGCGGCGGATGTCCTTCTTGAG
CCGGTGACGGCTGCGCTGGAGTGACGCTTGGGCGAGGAGGCACTGGCGCTCGGCTTGACAGAGAAATGCACATC
ACTGATGCGGAGGGCCTCGGCCTGGGCCCGCGCCTTCAGGAGGAAAGTTTGGGTTTGGGTCCCTCTTCTTGGG
35 CCCATACAGCTCACGCTCCCTCTCTTTTGTTCGAAGGCTGCAATGAGCCGCGAGTCCAGGATGTTCTCTCTCGG
CTCCAAGTGCTGT

SEQ ID NO:54

40 >GL3_37C_1_M13R Direction: anti-sense

ACAGCACTTGGGAGCCCCGAGGAGAACATCCTGGACTCGCGGCTCATTGCAGCCTTCGAACAAAAGGAGAGGGAGC
GTGAGCTGTATGGGCCCAAGAAGAGGGGACCCAAACCCAAAACCTTCTCTCTGAAGGCGCGGGCCCAGGCCGAGG
CCCTCCGCATCAGTGATGTGCATTTCTCTGTCAAGCCGAGCGCCAGTGCTCTCTCGCCCAAGCTGCACTCCAGCG
CAGCCGTGCACCGGCTCAAGAAGGACATCCGCCGCTGCCACCGTATGTCCCGCGTCCCTTGCCCCCGCCGGACC
45 CGCAGGGGGGCGAGCCCCGACTGCGCCCGCCATTTCGCCCTTCTCGGAGACGGTGCGCATCATCAACCGCAAGG
TGAAGCCGCGGGAGCCCAAGCGGAACCGCATCATCCTGAACCTGAAGGTGATCGACAAGGGCGCTGGCGGCGGGA
GGCGCCGGGCGAGGGGGCCGGGNCCTGGCCCGCCCCCAAAGTCCCTCGCGGAAACCCGCGGTAAATAGGCAAAG
AAGCAAGGAAGGTGAGGCGAGAGGCGGTCCCTGCGTCCGNAGAAACCCAGCACACCTGGACGGCCACCGATGCCA
CNGGCCTGGGGGACTCCTCTGGTACACGGTTCGGCACATGCCGG

SEQ ID NO:55

50 >GL3_37C_2_M13R Direction: anti-sense

ACAGCACTTGGGAGCCCCGAGGAGAACATCCTGGACTCGCGGCTCATTGCAGCCTTCGAACAAAAGGAGAGGGAGC
GTGAGCTGTATGGGCCCAAGAAGAGGGGACCCAAACCCAAAACCTTCTCTCTGAAGGCGCGGGCCCAGGCCGAGG
CCCTCCGCATCAGTGATGTGCATTTCTCTGTCAAGCCGAGCGCCAGTGCTCTCTCGCCCAAGCTGCACTCCAGCG
55 CAGCCGTGCACCGGCTCAAGAAGGACATCCGCCGCTGCCACCGTATGTCCCGCGTCCCTTGCCCCCGCCGGACC
CGCAGGGGGGCGAGCCCCGACTGCGCCCGCCATTATCAGCCCTTCTCGGAGACGGTGCGCATCATCAACCGCAA
GGTGAAGCCGCGGGAGCCCAAGCGGAANCGCATCATCCATGAACCTGAA

CAGGATGGGGCGGAGGCCTGGGCCCAGCTGGCAGGTCCCCTGGCATCGCAGGCACCTGTGGAGAGGGCCTG
 GACCCAGATCTCCACACCCGTGCTTGCTCAAAGGGAAGGACAACAGCGGGCCCCGGGAGCTAACCCAAG
 CTGCAGGTCCCAGCAAGCTGAGGTTTGGGAGGGTGGGGTTGTCACTGGTGATTTTCTCCAGGGGGCTGG
 TGAGTGGGCAGTTTGGTTTCTTGCCCCCTTCTGTTCTTTCCAGTTGTTGGGCCATCTGGTCCCCACCA
 5 CCGCCACCCCTATGGGGGAGACCTCCCTCCCCACGGGTACCCCTAAAGCCCACAACCTCTCTGAGCCTCCC
 TGGCCTGAAAGGGGATGCAGGCTTCAGGAGGCAAGAAGCTGGGCCCCCTGGGGGTGGCTGGGGAGAGGGAA
 TGCATTTCCCTTGCCACAGGTGGTCTGCTTCTGCTGGCCTGAGCTCCAAGTGGAGCAGCCCCGGGCCAGCC
 TTGGTGCATGAAGAGGCACCAGGCACACCGCCTTGAGGTGGGCAGTGGCCATGGGGGCCGAGTGGATGG
 GACCGAGGGTGAGTGGAGCCTCCTTCTCCCTCTCTAGTACCCCCGCCTCCACACACTTGCACGGATC
 10 GGCTCCCTTGGGAGATCAGCCTCCATGGGCCCCCTCGTCCACCCCTTGCTGCTTTCCATTTGCCTAATTAC
 CAAGCAGAAGTTGCAATCTGGTTTGTCTTTATTTTGTATGTGAAATAACCCCCAAAGCCCAATCTCCTCC
 TACGTTCAATATTGGTTGGGGCATCCGTCTATCTCCCTTAAAGTGCGCCCCCTCCCCACCCAAGTATCATA
 GGAAACCGGTGAGGTCTGGTGTCTCTGGTTTGAGACGGTAAGTTGGGACCCATCCCTGTCTGGGTGCCCA
 CTCTGACCTTTAGTTTGGCCTTCTGTGAAATGGGTGTATTGGGTAGCAAGCCCTCTTCAGAAAGCGCTGC
 15 TGGTGTGAGAGCAGCTGCCCAGTACCAGGTGGGGGTCAAGGTTGCTGGTACTGGGGCCCCCAGCTGCCC
 ACAACCCCTCTTTGTTCTCACCCCTGCAAAGGGGTCAAGGTCAAATGAGCCTCATCCTTCCTATGATCTG
 GGAAGAGGTGATGATCAAGTCCCCAACTTCAGTGTGAGGTGGACAGAGTTGGGGGGATGGCCCCCTTTTG
 AAGAGGTGAAATGGTTTTGGAGAAGCGCAGCTGCTTCACTGGGGGAATGCGGCAGGGACTGGGGCCAG
 GATGCTTTGGCCATGGGGAAAAGCCCTTTAAAGGCAGGGGCCAGGCCCTGGAGCCAGCACAAAGACTGG
 20 CCTCGAGCCCCTGAGCCAGGAGGTCTGGAGGAGAGCCAGGCCGGTGGGCCCCGCCAAGGCTGGAGGGTC
 AGCCCCAACAGGGAGCTGGGTGGCCAGGGGGCTGGACTGCTACCAGCCTCTCTGGCCTATGGGGACCCA
 AGAGGCACACATCCCCCTTTGCCACTCTTCTGTGTCATTTTGTGTTTGGTTTGTGGTGGTGGTGGTGGT
 TTTTCTTTTGTGGT
 TTTTCTTTTGTGGT
 25 TTTTCTTTTGTGGT
 TCCCTATCTGCATGCTCCTCTCATATAGAAAGCCCAGAATTAGGGGCTAGAACTCCAGGAGAGGGTCTCCC
 CGACTCATCTCTTGCTGACGGTCACCAGGATGCAGAAATAGGGAGATGGTTAGTGGGGGCCAAAGATGCC
 CCCTCCCAGGCCTTCGTGGTTCCCTCCTCCGCCCCCTGCAATCTTTGGAGGAGTCAGTGCCTCACTCCAG
 CAGTGAGTGCTACTGTATGCAGGTAGTCAGCCAGGCAAAGAGAGACTAACGGTCTCATGGGGGAACCCCT
 30 CTTGCGGGAGGCCCGGTAGCTGGAGCGAAGCGTTCGGCTGCCCTTGCTGCTGGGTGGAGTGGAGAGGGA
 GACTTCTTTTGTGGT
 ATTTTGGAGACGTTCAATTAATGAATTGTGCACGAATGAATTCTATATATATAAAATATACATATATAGC
 TCTATATTTGGGGAGGGGCACGTGCTCTTTTCTCTCATTTTAAATGAAGTGTGTTGCCTTTGTAT
 35 GTGGTTCAACCATCCAGCTCCCAGCTGGCTAAACTTTGCCTCCAGTGGTCAAAGATGGGAAAAGAGTGGG
 GTTGGCAGGAGATGGAAAACGGAGGTGCCGCCCCAGCATGGGGGGCAGGTCCCCAGTCCACCTGCCCC
 TCCCCCTGTGGAGAAGACGCTTAGTTGGGGGTGTGGGTTTGGGCTCCATTCTGGATTGGGCGGTTCGGG
 GGAGGGGTGGGTCTGTGCCGATTACTCTGTCTGTACGTTTGTCTGCTGCTCTTCAATATTGTATCAAC
 GCCAGGAAAGGGGGGTGAAAAGCCTCTTTTACCCCCCAAATAAATTGTACATTCCGAAGCTGA

40 SEQ ID NO:58

>gi|10140849|ref|NP_055107.1| chromobox homolog 6 [Homo sapiens]
 MELSAVGERVFAAESIIRIRKGRIEYLVKWKWAIKYSTWEPEENILDSRLIAAFEQKERERELYGPK
 KRGPKEPTFLKARAQAEALRISDVHFSVKPSASASSPKLHSSAAVHRLKKDIRRCHMSRRPLRPDPQ
 45 GGSPGLRPPISPFSETVRIINRKVKPREPKRNRIILNLKVIDKGAGGGGAGQGAGALARPKVPSRNRVIG
 KSKKFSSESVLRQIRHMKFGAFALYKPPAPLVPSPGKAEASAPGPGLLLAAPAAPYDARSSGSSGCPS
 PTPQSSDPDDTPPKLLPETVSPSPSWREPEVLDLSLPESAATSKRAPPEVTAAGPAPPTAPEPAGAS
 SEPEAGDWRPEMSPCSNVVTDVTSNLLTVTIKEFCNPEDFEKVAAGVAGAAGGGGSIGASK

50 SEQ ID NO:59

Cytochrome C Oxidase subunit 1
 >GL2_226_2_1_M13R Direction: sense
 ACTTCTCGGCCTATCCGGAATGCCCGACGTTACTCGGACTACCCCGATGCATACACCACATGAAACATCCTAT
 CATCTGTAGGCTCATTCATTTCTTAACAGCAG

55 SEQ ID NO:60

>GL2_226_2_2_M13F Direction: anti-sense
 CTGCTGTTAGAGAAATGAATGAGCCTACAGATGATAGGATGTTTCATGTGGTGTATGCATCGGGGTAGTCCGAGT
 AACGTCGGGGCATTCGGGATAGGCCGAGAAAGT

SEQ ID NO:61

>GL2_226_2_2_M13R Direction: sense

ACTTTCTCGGCTATCCGGAATGCCCCGACGTTACTCGGACTAACCCGATGCATACACCACATGAAACATCCTATN
ATACTGTAGGCTCCATTCAATTTCTCTNACAGCAGA

SEQ ID NO:62

>GL2_226_2_4_M13F Direction: anti-sense

CTGCTGTTAGAAGAAATGAATGAGCCTTACAGATGATAGGATGTTTCATGTGGTGTATGCATCGGGGTAGTCCGA
GTAACGTCGGGGCATTCGGGATAGGCCGAGAAAGT

SEQ ID NO:63

5905-7446

/gene="COX1"

/codon_start=1

/transl_table=2

/product="cytochrome c oxidase subunit I"

```
5881 agccatttta cctcaccccc actgatgttc gccgaccggt gactattctc tacaaaccac
5941 aaagacattg gaacactata cctattattc ggcgcgatgag ctggagtcct aggcacagct
6001 ctaagcctcc ttattcgagc cgagctgggc cagccaggca accttctagg taacgaccac
6061 atctacaacg ttatcgtcac agcccatgca tttgtaataa tcttcttcat agtaataccc
6121 atcataatcg gaggccttgg caactgacta gttccccctaa taatcggtgc ccccgatatg
6181 gcgtttcccc gcataaacia cataagcttc tgactcttac ctccctctct cctactcctg
6241 ctgcgatctg ctatagtggg ggccggagca ggaacagggt gaacagtcta ccctccctta
6301 gcagggaact actcccaccc tggagcctcc gtagacctaa ccatcttctc cttacaccta
6361 gcagggtgtct cctctatctt agggggccatc aatttcatca caacaattat caatataaaa
6421 cccctgcca taaccaata ccaaacgccc ctcttcgtct gatccgtcct aatcacagca
6481 gtctacttct tcctatctct cccagtccta gctgctggca tcaactatact actaacagac
6541 cgcaacctca acaccacctt ctctgacccc gccggaggag gagaccccat tctataccaa
6601 cacctattct gatttttctcg tcaccctgaa gtttatattc ttatcctacc aggettccga
6661 ataattctccc atattgtaac ttactactcc ggaaaaaaag aaccatttgg atacataggt
6721 atgggtctgag ctatgatatc aattggcttc ctagggttta tcgtgtgagc acaccatata
6781 tttacagtag gaatagacgt agacacacga gcatatttca cctccgctac cataatcatc
6841 gctatcccca ccggcgtcaa agtatttagc tgactcgcca cactccacgg aagcaatatg
6901 aatgatctg ctgcagtgct ctgagcccta ggattcatct ttcttttcac cgtagggtggc
6961 ctgactggca ttgtatttag aaactcatca ctagacatcg tactacacga cagtactac
7021 gttgtagctc acttccacta tgtcctatca ataggagctg tatttgccat cataggaggc
7081 ttcattcact gatttcccct attctcaggc tacaccctag accaaacctc cgccaaaatc
7141 catttccact tcatattcat cggcgtaaat ctaactttct tcccacaaca ctttctcggc
7201 ctatccggaa tgccccgacg ttactcggac taccctgatg catacaccac atgaacatc
7261 ctatcatctg taggctcatt catttctcta acagcagtaa tattaataat tttcatgatt
7321 tgagaagcct tcgcttcgaa gcgaaaagtc ctaatagtag aagaacctc cataaacctg
7381 gaggtagctat atggatgccc ccaccctac cacacattcg aagaacctg atacataaaa
7441 tctagacaaa
```

SEQ ID NO:64

```
>gi|17981855|ref|NP_536845.1| cytochrome c oxidase subunit I [Homo sapiens]
MFADRWLFSTNHKDIGTLYLLFGAWAGVLGTALSLLIRAE LGQPGNLLGNDHIYNVIVTAHAFVMIFFMV
MPIMIGGFNGWLVPLMIGAPDMAFPRMNNMSFWLLPPSLLLLLASAMVEAGAGTGWTVYPPLAGNYSHPG
ASVDLTIFSLHLAGVSSILGAINFITTIINMKPPAMTQYQTPLFVWSVLITAVLLLLSLPVLAAAGITMLL
TDRNLNTFFDPAGGGDPILYQHLFWFFGHPEVYILILPGFGMISHIVTYYS GKKEPFGYMGMVWAMMSI
GFLGFIVWAHMFVGMVDVTRAYFTSATMIIAIP TGVKVFSWLATLHGSNMKWSAAVLWALGFIFLFTV
GGLTGIVLANSSLDIVLHDTYYVVAHFHYVLSMGAVFAIMGGFIHWFP LFSGYTL DQTYAKIHFTIMFIG
VNLTFFPQHFLGLSGMPRRYSYDPDAYTTWNILSSVGSFISLTAVMLMIFMIWEAFASKRKVLMVEEPSM
NLEWLYGCPPPYHTFEFPVYMK
```

SEQ ID NO:65

Delta-5 Fatty Acid Desaturase (FADS1; FADS5)

>GL1_10_3_G3F1 Direction: sense

GGCCTGAAC TGNACTGAGCAGCACCGCACANAGGAGGAAGGGCAAAAAGGACGTCCCAAAGACCCAAAGGGTGAG
CCAGNCTGNACCATCCAGCAGCAAGATGTGCAGACAGGTACAGNAGGAAGAAACATGGTTGGCCTTCATGAGCC

CCATCCGCTCCACTGTGGNCCGCANACTCCCGGAATCATCTGTCAGCTCTTTATTCTTGGTGGGNTCAAAGCTG
GGCTGNTCTGAGACAGTTCTCCAATCAGGAGAGAGTTTCATATACTTCTTCACAAGGCCCTTGTGGATGTGGAAG
GCCACAAAGGGATCCGTGGNATCCTGNCCGGNGTAGTGGCTGATGANCCGGGAGCCCCCTGNATGCCGGCGGAGT
GAACTCGCTGATGTTGTACACCTTACGGTCGATCACTANCCACCGCTCCT

SEQ ID NO:66

>GL1_10_3_G3R1 Direction: anti-sense

CAGACGCTCAAGGATGCGAGGAGCGGTGGCTAGTATCGACCGNTAAGGTGTACAACATCAGCGAGTTCACCCGC
CGGCATCCAGGGGGCTCCCGGGTTCATCAGCCACTACGCCGGGAGGATGCCACGGATCCCTTTGTGGCCTTCCAC
ATCAACAAGGGCCTTGCTGAAGAAGTATATGAACTCTCTCCTGATTGGAGAACTGTCTCCAGAGCAGCCAGCTT
TGAGCCACCAAGAATAAAGAGCTGACAGATGAGTTCCGGGAGCTGCGGGCCACAGTGGAGCGGATGGGGCTCAT
GAAGGCCAACCATGTCTTCTTCTGCTGTACCTGCTGCACATCTTGCTGCTGGATGGTGCAGCCTGGCTCACCTT
TTGGTCTTTGGGACGTCCTTTTGGCCCTTCTCCTCTGTGCGGTGCTGCTCAGTGCAGTTCAGGCC

SEQ ID NO:67

delta5-Fatty acid desaturase>

ATGGCCCCCGACCCGTTGGCCCGGAGACCGCGGCTCAGGGACTTACCCC
GCGCTACTTCACCTGGGACGAGGTGGCCAGCGCTCAGGGTGCAGGAGC
GTTGGCTAGTGTATCGACCGTAAGGTGTACAACATCAGCGAGTTCACCCGC
CGGCATCCAGGGGGCTCCCGGGTTCATCAGCCACTACGCCGGGAGGATGC
CACGGATCCCTTTGTGGCCTTCCACATCAACAAGGGCCTTGTGAAGAAGT
ATATGAACTCTCTCCTGATTGGAGAACTGTCTCCAGAGCAGCCAGCTTT
GAGCCACCAAGAATAAAGAGCTGACAGATGAGTTCCGGGAGCTGCGGGC
CACAGTGGAGCGGATGGGGCTCATGAAGGCCAACCATGTCTTCTTCTGCT
TGTACCTGCTGCACATCTTGTGCTGCTGGATGGTGCAGCCTGGCTCACCTT
TGGGTCTTTGGGACGTCCTTTTGGCCCTTCTCCTCTGTGCGGTGCTGCT
CAGTGCAGTTTCAAGCCAGGCTGGCTGGCTGCAGCATGACTTTGGACACC
TGTCGGTCTTCAGCACCTCAAAGTGAACCATCTGCTACACCATTTTGTG
ATTGGCCACCTGAAGGGGGCCCCCGCCAGTTGGTGAACCATGCACTT
CCAGCACCATGCCAAGCCCAACTGCTTCCGCAAAGACCCAGACATCAACA
TGCATCCCTTCTTCTTGGCCTTGGGGAAGATCCTCTCTGTGGAGCTTGGG
AAACAGAAAGAAAATTATATGCCGTACAACCACCAGCACAAATACTTCTT
CCTAATTGGGCCCCCGGCCCTTGTGCCTCTCTACTTCCAGTGGTATATTT
TCTATTTTGTATCCAGCGAAAGAAGTGGGTGGACTTGGCCTGGATGATT
ACCTTCTACGTCCGCTTCTTCTTCACTTATGTGCCACTATTGGGGCTGAA
AGCCTTCTTGGGCTTTTCTTTCATAGTCAGGTTCTTGGAAGCAACTGGT
TTGTGTGGGTGACACAGATGAACCATATTCCCATGCACATTGATCATGAC
CGGAACATGGACTGGGTTTCCACCCAGCTCCAGGCCACATGCAATGTCCA
CAAGTCTGCCTTCAATGACTGGTTTCAAGTGGACACCTCAACTTCCAGATTG
AGCACCATCTTTTCCCACGATGCCTCGACACAATTACCACAAAGTGGCT
CCCCTGGTGCAGTCTTGTGTGCCAAGCATGGCATAGAGTACCAGTCCAA
GCCCCGTGCTGCAGCCTTCGCCGACATCATCCACTACTAAAGGAGTCAG
GGCAGCTCTGGCTAGATGCCTATCTTCACCAATAA

SEQ ID NO:68

>gi|14141179|ref|NM_013402.3| Homo sapiens fatty acid desaturase 1 (FADS1), mRNA

TCCACTCCTGGAGCCCGCGGACCCGAGCACGCGCCTGACAGCCCTGCTGGCCCGGCGCGCGGCTCGC
CAGGCCAGCTATGGCCCCCGACCCGGTGGCCGCGGAGACCGCGGCTCAGGGACCTACCCCGCGCTACTTC
ACCTGGGACGAGGTGGCCAGCGCTCAGGGTGCAGGAGCGGTGGCTAGTGATCGACCGTAAGGTGTACA
ACATCAGCGAGTTACCCGCGCGCATCCAGGGGGCTCCCGGGTTCATCAGCCACTACGCCGGGAGGATGC
CACGGATCCCTTTGTGGCCTTCCACATCAACAAGGGCCTTGTGAAGAAGTATATGAACTCTCTCCTGATT
GGAGAACTGTCTCCAGAGCAGCCAGCTTTGAGCCACCAAGAATAAAGAGCTGACAGATGAGTTCGGG
AGCTGCGGGCCACAGTGGAGCGGATGGGGCTCATGAAGGCCAACCATGTCTTCTTCTGCTGTACCTGCT
GCACATCTTGCTGCTGGATGGTGCAGCCTGGCTACCCCTTTGGGTCTTTGGGACGTCCTTTTGGCCCTT
CTCCTCTGTGCGGTGCTGCTCAGTGCAGTTTCAAGCCAGGCTGGCTGGCTGCAGCATGACTTTGGGCACC
TGTCGGTCTTCAGCACCTCAAAGTGAACCATCTGCTACATCATTTTGTGATTGGCCACCTGAAGGGGGC
CCCCGCCAGTTGGTGAACCATGCACTTCCAGCACCATGCCAAGCCCAACTGCTTCCGCAAAGACCCA
GACATCAACATGCATCCCTTCTTCTTGGCCTTGGGGAAGATCCTCTCTGTGGAGCTTGGGAAACAGAAGA
AAAAATATATGCCGTACAACCACCAGCACAAATACTTCTTCTAATTGGGCCCCCGAGCTTGTGCTGCTCT

CTACTTCCAGTGGTATATTTTCTATTTTGTATCCAGCGAAAGAAGTGGGTGGACTTGGCCTGGATGATT
 ACCTTCTACGTCCGCTTCTTCTCACTTATGTGCCACTATTGGGGCTGAAAGCCTTCCTGGGCCTTTTCT
 TCATAGTCAGGTTCTTGGAAGCAACTGGTTTGTGTGGGTGACACAGATGAACCATATTCCTATGCACAT
 5 TGATCATGACCGGAACATGGACTGGGTTTCCACCCAGCTCCAGGCCACATGCAATGTCCACAAGTCTGCC
 TTCAATGACTGGTTCAGTGGACACCTCAACTTCCAGATTGAGCACCATCTTTTTCCCACGATGCCTCGAC
 ACAATTACCACAAAGTGGCTCCCCTGGTGCAGTCTTGTGTGCCAAGCATGGCATAGAGTACCAGTCCAA
 GCCCTGTCTGTCAGCCTTCGCGGACATCATCCACTCACTAAAGGAGTCAGGGCAGCTCTGGCTAGATGCC
 TATCTTCCACCAATAACAACAGCCACCCTGCCCAGTCTGGAAGAGAGGAGGAAGACTCTGGAGCCAAGGC
 10 AGAGGGGAGCTTGAGGGACAATGCCACTATAGTTTAATACTCAGAGGGGGTGGGTTTGGGGACATAAAG
 CCTCTGACTCAAACCTCCTCCCTTTTATCTTCTAGCCACAGTTCTAAGACCCAAAGTGGGGGGTGGACACA
 GAACTCCCTAGGAGGGAAGGAGCTGTTGGGGCAGGGGTGTAATTTATTTCTTTTCTAGTTTGGCACAT
 GCAGGTAGTTGGTGAACAGAGAGAACCAGGAGGGTAACAGAAGAGGAGGGACCTACTGAACCCAGAGTCA
 GGAAGAGATTTAACACTAAAATTCCACTCATGCCGGGCGTGGTGGCACGCGCTGTAATCCAGATACCC
 15 AGGAGGCTGAGGCAGGAGAATCGCTTGAACCGGGGAGGTGGAGGTTGCAGTGAGCTGAGATCACGCCATT
 GTACTCCAGCCTGGGCGACAGAGCAAGACTCCATTTCAAAAAAAAAAAAAAAAAATCCACTCATATAAAAGG
 TGAGCTCAGCTCACTGGTCCATTTCTCAGTGGCTTCTCCATCCTCATTTGCAAACCTCAGAGGGATAAGG
 CAGTTGAACCTGATGAGCAAGAATTATAACAGCAAGGAAACATTAATGCTTAGAATTCAGATCCAGCA
 CAACTCAGTCTGTGGGAGCTCAGCTCGCTGCCAGGGATAGGTATGACCTATGTCTGCCCTTAGGCTGCTG
 20 GGAGATGCCATTCTCCAGTTTTCAGAAGCAGGCAGGGCAAAGGTCAAGACTGTGGTATTGGGGTCTTTTGG
 CTCTGAAGGATCCTGGAACCACTGATTTTGGTTTATTTCCCTCCAGGGTCTAAAGAGAACAAAGAGGTGCTA
 GCTCTTACCAAAACAGATGGTAGAGAGAGTTGCTGGCTATTTAAAAAGCTCTTTCATCTTTTAAATTCACC
 TCTTCTTTTCACTCTTTTAAACCACTCCTCAGGAACAGAACACTTCTAGGACTGGGGGTCTTTTAGCTCCA
 TAAGCAAGTGAGCAGATGGGACAAGTTAGTCTTTTCTCCCTAGAAACAAAGGGGATGCCAGTGGTTTCC
 25 CTTTGGCTTCCCAACCTAAAATTTCAAGTTTAAATAAATAGCAATTAGCAGAAGTGACCAAATGGGAGAT
 AATTATCAGTCATGAGGAAAGACACAGATTTTCGGTCATAAAGAATGTAAGGGCTATAAGTAGAACTTTC
 TATAACCTAAATGATGTTATAGAATTATTTTGTAGCAGGAGCAGAAAGATTAAATATGATCACTTCATAC
 TTCTAAATCAGAAATAGGAAGATTAAACCACAGAACAGTTTGTGATTTCTATTGCTGTAGCTAGGTATC
 TTACTCTGTCCACTCTTGTTCAGTATCTAACTCTTCTGGAACCAAATAGGCTTTAGAAAGAGATTATCC
 30 TATATTCCCTATCAGTATAATACTAAATGTAACTTTAAATCATCTGGTTTAAAGATAAACAGTTTA
 GCCCATCTCTCCAGAGAGCAAACATAGGAATATGACTCAGGAGCCTCCTAGGCTTATCATCAGCCCTCA
 CACCCGCTTCCCCCTCCAACCCACAGCCTTTGCTTCCAGGTGGCAGGATTACTACTTTGCTTTCAGCA
 GCATCTACTCTAGGCATATTGATCATTTTACACTGGGAGAAGAGAACCTCAAACCTAGGAGGAAAAGAC
 AGAGCCTCCACTTAGTTTGGGAGGGGATGGCAGACAGTCAAGGAGATGAGCGTCTAAGGCATGTTGGG
 35 ATAGGGTCAGATGCACCACCCATGGAGAGGTTTGTCAACACAAAGACATGGAAGGTTAGAGGTTTGTCAA
 CAAAAAGACATGGAAGGTTAGGTTTGTCAACACAAAGACATGGAAGATTAGAGGTTTGTCAACACAAAGA
 CACAGGAAGAATGGGCTGCAGAAGATTAGATGTTTCCATTTGGGCACATTTTACTTAGCTGGAGAACT
 AGGTTTAAACAGCCTGGGTAGGAAAATTAGAAGCAAGCTGGATGCAGTGGCTCATGCCTGTAATCCCAA
 CACTTTTGGGAGGTCAGGCAGGAGGATCACTTGGGCCAGGAGGTCAAGCCTGCAGCGAGCTGAGATCA
 40 CACCACTGCACTCCAGCCTGGGGTGATAGAACAAGACCCTGTCTCAAAAAAAAAAAAAAAAAACAAAAAC
 TTAGAATTGAGGAGTTGTACCTCCATTGGCTTCTCACTCCAAAATAGGTGCTGATCCTTCTTATTCCTA
 TTCTTTGCCACCTTTTGGGTGTGGTGTACCAGCCTGTTTAGCCAAGTAGCTTTGGGCATAGGCTGCCCA
 ATCTGAGCAAACACAGTGAGGCTCTATTGAGCCAAGACCAAGTCCTCAAAGCACCTGAACCACTGTGGC
 CTTCTCAGCCTACAGCAGTGTGGTCTCTTACATGGCCACAAAGGGACACACAGTGACAAAAGGCTCGGAA
 45 TGTTACAATGGTAAAATGAGTGATCTCAAATCCACTGACAGATATAAAATAGGCTTAGAGAGGAAAAGCT
 GCCTCTGGTCAAGTAGATCATGGCAGCATGAATCCAACTCACTTTTTTACAACCTCCAACCTTCTATGTTT
 ATCTTTGTTACTTTCACTTTTTTACAACCTGGCCAGAGGCATTTTTTAAATCAGGCCCAATATCAGTATT
 CTTTTTGTGTGTGCCAATTTTGTATCACATCCCTATGAAGTTGAAAAATAAAGTTAATTTTGACCAAAA
 AAAAAAAAAAAG

50 SEQ ID NO:69
 >gi|11181775|ref|NP_037534.2| fatty acid desaturase 1; delta-5 fatty acid
 desaturase; linoleoyl-CoA desaturase (delta-6-desaturase)-like 1; delta-5
 desaturase [Homo sapiens]
 MAPDPVAAETAAGQPTPRYFTWDEVAQRSGCEERWLVIDRKVYNISEFTRRHPPGGSRVISHYAGQDATDP
 55 FVAFHINKGLVKKYMNSLLIGELSPEQPSFEPTKNKELTDEFRELRAFTERMGLMKANHVFFLLYLLHIL
 LLDGAAWLLTWVFGTSFLPFLLCVLLSAVQAQAGWLQHDGHLVSFSTSKWNHLLHHFVIGHLKGAPAS
 WVNHHMFQHHAKPNCFRKDPDINMHPFFALGKILSVELGKQKKYMPYNHQHKYFFLIGPPALLPLYFQ
 WYIFYFVIQRKKWVDLAWMITFYVRFFLTIVVPLGLKAFGLGFFIVRFLESNWFVWVTQMNHIMHIDHD
 RNMDWVSTQLQATCNVHKSANFNDWFSGHLNFQIEHHLFPTMPRHNYHKVAPLVQSLCAKHGIEYQSKPLL
 60 SAFADIHSLKESGQLWLDAYLHQ

SEQ ID NO:70

Dynactin 6, p27 subunit

>gi|18426895|ref|NM_006571.2| Homo sapiens novel RGD-containing protein (WS-3), mRNA

5 CAACCCTGCCAGGCTCTCCAATCGCATGTGGAATTATCGCTCTACCCAGGCGGTGGTGTGCGATCTACGTT
CCAATTGGGGCCGTACCATGGCGGAGAAGACTCAAAGAGTGTGAAGATTGCTCCTGGAGCAGTTGTATG
TGTAGAAAGTGAAATCAGAGGAGATGTAACATATCGGACCTCGGACAGTGATCCACCCTAAAGCAAGAATT
ATTGCGGAAGCCGGGCCAATAGTGATTGGCGAAGGGAACCTAATAGAAGAACAGGCCCTTATCATAAATG
10 CTTACCCAGATAATATCACTCCTGACACTGAAGATCCAGAACCAAAACCTATGATCATTGGCACCATAA
TGTGTTTTGAAGTTGGCTGTTATTCCCAAGCCATGAAGATGGGAGATAATAATGTCATTGAATCAAAAGCA
TATGTAGGCAGAAATGTAATATTGACAAGTGGCTGCATCATTGGGGCTTGTTGCAACCTAAATACATTTG
AAGTCATCCCTGAGAATACGGTGATCTATGGTGCAGACTGCCTTCGTCGGGTGCAGACTGAGCGACCGCA
GCCCCAGACACTACAGCTGGATTTCTTGATGAAAATCTTGCCAAATTACCACCACCTAAAGAAGACTATG
15 AAAGGAAGCTCAACTCCAGTAAAGAACTAAGAACAGTGTATAACATGAAGATAACATTTTGTCTTTGACC
ACTGTCTTTTGAATGGGCCACAGTGTGTTTGTACTCTTAACAACCTCACAGAATAATACATGTTCACTTT
ATTTTGTAAAATTGGGTTGAGAGGAACTAATGGAGTTTCATTGTAAGTGTCTTTGTAAATTTATATAAA
TGTATTATTTTCTATATCCTTGGTCTTTTCTGATAATTTACAGATTTAGCTTTTCTTTGTTATATAA
ACTGCTAGCCACAAATTTTAGTTATGTAAAAGGCTACCCTTGACAAGAAAAGACATACTGTCATGTATTT
ATATTCTAGCATAGACTAACTGAATAAAAATGCTGATAACAGGACCTTTAAAAA

SEQ ID NO:71

>gi|5730116|ref|NP_006562.1| novel RGD-containing protein [Homo sapiens]

MAEKTQKSVKIAPGAVVCVESEIRGDVTIGPRTVIHPKARIIEAGPIVIGEGNLIEEQALIINAYPDNI
25 TPDTEDEPEPKPMIIGTNNVFEVGCYSQAMKMGDNNVIESKAYVGRNVILTSGCIIGACCNLNTFEVIPEN
TVIYGADCLRRVQTERPQPQTLQLDFLMKILPNYHHLKKTMKGSSTPVKN

SEQ ID NO:72

Elongation factor 1 alpha (EF1 α)

>GL1_34_LOW_3_G3F1 Direction: sense

30 ACTGTNCTGCTTGGCCGAGTGGAGACTGGTGTCTCAAACCCGGATATGGTGGTGCACCTTTGCTCCAGCTCAAC
GTTACAACGGAAGTAAATCTGTGCAATGCACCATGAAGCTTTGAGTGAAGCTCTTCTGGGGACAATGTGGGC
TTCAATGTCAAGAATGTGTCTGTCAAGGATGTTCTGTCGTGGCAACGTNGCTGGTGACAGCAAAATGACCCACCA
ATGGAAGCAGCTGGCTTCACTGCTCAGGTGATTATCCTGAACCATCCAGGCCAAATAAGCGCC

SEQ ID NO:73

>GL1_34_LOW_3_G3R1 Direction: anti-sense

35 GGCGCTTATTTGGCCTGGATGGTTCAGGATAATCACCTGAGCAGTGAAGCCAGCTGCTTCCATTGGTGGGTGCTATT
TTTGCTGTCCAGCAACGTTGCCACGACGAACATCCTTGACAGACACATTCTTGACATTGAAGCCACATTGTC
CCCAGGAAGAGCTTCACTCAAAGCTTCATGGTGCATTTGACAGATTTTACTTCCGTTGTAACGTTGACTGGAGC
40 AAAGGTGACCACCATAACCGGGTTTGAGAACACCAGTCTCCACTCGGCCAACAGGAACAGT

SEQ ID NO:74

>GL1_34_LOW_4_G3F1 Direction: sense

45 ACTGTTCTGTTGGCCGAGTGGAGACTGGTGTCTCAAACCCGGTATGGTGGTGCACCTTTGCTCCAGTCAACGTT
ACAACGGAAGTAAATCTGTGCAATGCACCATGAAGCTTTGAGTGAAGCTCTTCTGGGGACAATGTGGGCTTC
AATGTCAAGAATGTGTCTGTCAAGGATGTTCTGTCGTGGCAACGTTGCTGGTGACAGCAAAATGACCCACCAATG
GAAGCAGCTGGCTTCACTGCTCAGGTGATTATCCTGAACCATCCAGGCCAAATAAGCGCC

SEQ ID NO:75

>GL1_34_LOW_4_G3R1 Direction: anti-sense

50 GGCGCTTATTTGGNCTGGATGGTTCAGGATAATCACCTGAGCAGTGAAGCCAGCTGCTTCCATTGGTGGGTGCTATT
TNTGNTGTCACCAGCAACGTTGCCACGACGAACATCCTTGACAGACACATTCTTGACATTGAAGCCACATTGTC
CCCAGGAAGAGCTTCACTCAAAGCTTCATGGTGCATTTGACAGATTTTACTTCCGTTGTAACGTTGACTGGAGC
55 AAAGGTGACCACCATAACCGGGTTTGAGAACACCAGTCTCCACTCGGCCAACAGGAACAGT

SEQ ID NO:76

>gi|20127674|ref|NM_001402.3| Homo sapiens eukaryotic translation

elongation factor 1 alpha 1 (EEF1A1), mRNA

GGCACGAGGGTTTGGCCGAGAACACAGGTGTCGTGAAAACCTACCCCTAAAAGCCAAAATGGGAAAGGAA
60 AAGACTCATATCAACATTGTCGTCAATTGGACACGTAGATTCTGGGCAAGTCCACCCTACTGGCCATCTGA

TCTATAAATGCGGTGGCAATGACAAAAGAACCATTGAAAAATTTGAGAAGGAGGCTGCTGAGATGGGAAA
GGGCTCCTTCAAGTATGCCTGGGTCTTGGATAAACTGAAAGCTGAGCGTGAACGTGGTATCACCATTGAT
ATCTCCTTGTGGAAATTTGAGACCAGCAAGTACTATGTGACTATCATTGATGCCCCAGGACACAGAGACT
TTATCAAAAACATGATTACAGGGACATCTCAGGCTGACTGTGCTGTCTGATTGTTGCTGCTGGTGTGG
5 TGAATTTGAAGCTGGTATCTCCAAGAATGGGCAGACCCGAGAGCATGCCCTTCTGGCTTACACACTGGGT
GTGAAACAACATAATTGTCGGTGTAAACAAAATGGATTCCACTGAGCCACCCTACAGCCAGAAGAGATATG
AGGAAATTGTTAAGGAAGTCAGCACTTACATTAAGAAAATTGGCTACAACCCCGACACAGTAGCATTGT
GCCAATTTCTGGTTGGAATGGTGACAACATGCTGGAGCCAAGTGCTAACATGCCTTGGTTCAAGGGATGG
AAAGTCACCCGTAAGGATGGCAATGCCAGTGGAAACCACGCTGCTTGAGGCTCTGGACTGCATCCTACCAC
10 CAACTCGTCCAAGTACAAGCCCTTGCGCTGCCCTCTCCAGGATGTCTACAAAATTGGTGGTATTGGTAC
TGTTCTCTGTTGGCCGAGTGGAGACTGGTGTCTCAAACCCGGTATGGTGGTCACTTTTGCTCCAGTCAAC
GTTACAACGGAAGTAAAATCTGTGCAAAATGCACCATGAAGCTTTGAGTGAAGCTCTTCTGGGGACAATG
TGGGCTTCAATGTCAAGAATGTGTCTGTCAAGGATGTTCTGCTGGCAACGTTGCTGGTGCAGCAAAAA
TGACCCACCAATGGAAGCAGCTGGCTTCACTGCTCAGGTGATTATCCTGAACCATCCAGGCCAAATAAGC
15 CGCGCTATGCCCCCTGTATTGGATTGCCACACGGCTCACATTGCATGCAAGTTTGCTGAGCTGAAGGAAA
AGATTGATCGCCGTTCTGGTAAAAAGCTGGAAGATGGCCCTAAATCTTGAAGTCTGGTGTATGCTGCCAT
TGTTGATATGGTTCTTGGCAAGCCCATGTGTGTTGAGAGCTTCTCAGACTATCCACCTTTGGGTGCTTT
GCTGTTCTGTATATGAGACAGACAGTTGCGGTGGGTGTCTCAAAGCAGTGGACAAGAAGGCTGCTGGAG
20 CTGGCAAGGTCAACAAGTCTGCCAGAAAGCTCAGAAGGCTAAATGAATATTATCCCTAATACCTGCCAC
CCCCTCTTAATCAGTGGTGAAGAACGGTCTCAGAACTGTTTGTTCATTGGCCATTTAAGTTTAGTA
GTAAAAGACTGGTTAATGATAACAATGCATCGTAAACCTTCAGAAGGAAAGGAGAATGTTTTGTGGACC
ACTTTGGTTTTCTTTTTTTCGCTGTGGCAGTTTTTAAGTTATTAGTTTTTTAAATCAGTACTTTTTTAATGGA
AACAACTTGACCAAAAATTTGTACAGAATTTTGAGACCCATTAAAAAAGTTAAATGAGAAAAA
25 AA
AAAAAAAAAAAA

SEQ ID NO:77

>gi|4503471|ref|NP_001393.1| eukaryotic translation elongation factor 1
alpha 1 [Homo sapiens]
30 MGKEKTHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKLKAERERG
ITIDISLWKFETSKYYVTIIDAPGHRDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTREHALLA
YTLGVKQLIVGVNKM DST EPPYSQKRYEEIVKEVSTYIKKIGYNPDVAFVPISGWNGDNMLEPSANMPW
FKGWKVRTRKDGNASGTTLLLEALDCILPPTRPDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTF
APVNVTTTEVKSVMHHEALSEALPGDNVGFNVKNVSKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHP
35 QGISAGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPKPMCVESFSDYPP
LGRFAVRDMRQTVAVGVKAVDKKAAGAGKVTKSAQKAQKAK

SEQ ID NO:78

Elongation factor 4 gamma (ElF4γ)
40 >GL1_3_1_G3F1 Direction: sense
GTCCAACCTCAAAGATATGGGAGAGGATTTGGAGTGCCTCTGTCTAGATAATGAGGACAGTGGGACCTAGATTAGAC
CATGAACGAGCCAAGTCCTTAATGGATCAGT

SEQ ID NO:79

>GL1_3_1_G3R1 Direction: anti-sense
45 ACTGATCCATTAAGGACTTGGCTCGTTCATGGTCTAATCTAGGTCCCACTGGTCCTCATTATCTGACAGAGGCAC
TCCAAATCCTCTCCCATATCTTTGAGTTGGAC

SEQ ID NO:80

>GL1_3_2_G3R1 Direction: anti-sense
50 ACTGATCCATTAAGGACTTGGNTCGTTCATGGTCTAATCTANAGTCCCACTGGTCCTCATTATCTGACAGGAGAG
CACTCCAAATCCTCTCCCATATCTTTGAGTTGGAC

SEQ ID NO:81

>gi|4503538|ref|NM_001418.1| Homo sapiens eukaryotic translation initiation
factor 4 gamma, 2 (EIF4G2), mRNA
55 CAGCAGTGAGTCGGAGCTCTATGGAGGTGGCAGCGGGTACCGAGTGGCGGCTGCAGCAGCGACTCCTCTG
AGCTGAGTTTGAGGCCGTCCCCGACTCCTTCCCTCCCCCTTCCCTCCCCCTTTTTTTTGTTCCTGTTCCGTTCC
CTTTCCCTCCCTTCCCTATCCCCGACGACCGGATCCTGAGGAGGCAGCTGCGGTGGCAGCTGCTGAGTT
60 CTCGGTGAAGGTATTTCAATTTCTCTGTCCCCTCCCCCTCCCCACCCCATCTATTAATATTATCTTTTGA

5 AAATAAGCTTACTCCTGAAAAGTTTGACAAGCTATGCTTGAGCTCCTCAATGTGGGTGTAGAGTCTAA
CTCATCCTTAAAGGGGTCATACTGCTGATTGTGGACAAAGCCCTAGAAGAGCCAAAGTATAGCTCAGTGT
ATGCTCAGCTATGTCTGCGATTGGCAGAAGATGCACCAAACCTTTGATGGCCCAGCAGCAGAGGGTCAACC
AGGACAGAAGCAAAGCACCACATTAGACGCCTCCTAATTTCCAAATTACAAGATGAATTTGAAAACCGA
10 ACTAGAAATGTTGATGTCTATGATAAGCGTGAAAATCCCCTCCTCCCCGAGGAGGAGGAACAGAGAGCCA
TTGCTAAGATCAAGATGTTGGGAAACATCAAATTCATTGGAGAGCTTGGCAAGCTTGATCTTATTCACGA
ATCTATCCTTCATAAGTGCATCAAAACAC'TTTTGGAAAAGAAGAAGAGAGTCCAAC'TCAAAGATATGGGA
GAGGATTTGGAGTGCCTCTGTGAGATAATGAGGACAGTGGGACCTAGATTAGACCATGAACGAGCCAAG'
CCTTAATGGATCAGTACTTTGCCCGAATGTGCTCCTTGATGTTAAGTAAGGAATTTGCCAGCAAGGATTCG
TTTCTGTGAGGATACCGTAGATTGGCAGAACACCATTTGGGTCTCTCGCAAGGCTTTTCTTGACAAT
15 GGACCAAAGACGATCAATCAAATTCGTCAAGATGCAGTAAAAGATCTAGGGGTGTTTATTTCTGTCTCCTA
TGGCTCAAGGGATGAGAAGTGACTTCTTTCTGGAGGGACCGTTTATGCCACCCAGGATGAAAATGGATAG
GGACCCACTTTGGAGGACTTGTGATATGTTTGGACAAATGCCAGGTAGCGGAATTTGGTACTGGTCCAGGA
GTTATCCAGGATAGATTTTCAACCACCATGGGACGTCTATCGTTCAAATCAACTCTTCAATGGCCATGGGG
GACACATCATGCCTCCACACAATCGCAGTTTGGAGAGATGGGAGGCAAGTTTATGAAAAGCCAGGGGCT
20 AAGCCAGCTCTACCATAACCAGAGTCAGGGACTCTTATCCCAGCTGCAAGGACAGTCGAAGGATATGCCA
CCTCGGTTTTCTAAGAAAGGACAGCTTAAATGCAGATGAGATTAGCCTGAGGCCTGCTCAGTCGTTCCCTAA
TGAATAAAAAATCAAGTGCCAAAGCTTCAGCCCCAGATAAATATGATTCCTCTAGTGCACAACCACCACG
CACTCAAAACCACCTCTGGGACAGACACCTCAGCTTGGTCTCAAAACTAATCCACCACCTTATCCAGGAA
AAGCCTGCCAAGACCAGCAAAAAGGCCACCCTCAAAGGAAGAAGACTCCTTAAACTAATCGAACTGTTG
25 TGACTGAATATCTAAATAGTGGAAATGCAAATGAGGC'TGTCAATGGTGTAAAGAGAAATGAGGGCTCTCTAA
ACACTTTCTTCTTGAGATGTTAAGCAAAGTAATCATCCTGTCTACTAGATAGAAGCGATGAAGATAAAGAA
AAAGCAAGTTCTTTGATCAGTTTACTCAAACAGGAAGGGATAGCCACAAGTGACAACCTTCATGCAGGCTT
TCCTGAATGTATTGGACCAGTGTCCCAAAC'TGGAGGTTGACATCCC'TTTGGTGAAATCCTATTTAGCACA
GTTTGCAGCTCGTGCCATCATTTTCAGAGCTGGTGAGCATTTTCAGAACTAGCTCAACCAC'TAGAAAGTGGC
30 ACCCATTTTCTCTCTTCTTCTTGTCTTCAGCAGTTAGCTAAAT'TACAAGATCGAGAAATGGTTAACAG
AACTTTTCTCAACAAAGCAAGGTCAATATGCAGAAAATGCTCCCAGAAATTTGATCAGAATAAGGACCGCAT
TTTGGAGATTTTGGAAAGGAAGGACTAGT'TTCTTATTTCCACTCTCAAATTTGGAGAAGGAAGCTTTG
AAGCAAATAAAGTTGGATCCATCCCCCTCAAACCATATATAAATGGATTAAAGATAACATCTCTCCCAAAC
TTCATGTAGATAAAGGATTTGTGAACATCTTAAATGACTAGCTTCTTACAGTACATTTCTAGTGAAGTAA
35 CCCCCCAGCGATGAAACAGATTTCATCCTCTGCTCCTTCCAAAGAACAGTTAGAGCAGGAAAAACAATA
CTACTATCTTTCAAGCCAGTAATGCAGAAAATTTCTTCATGATCACGTTGATCTACAAGTCAGTGCCCTGT
ATGCTCTCCAGGTGCATGCTATAACAGCAACTTCCCAAAGGCATGTTACTTCGCTTTTTTGTGCACTT
CTATGACATGGAAATTAATTGAAGAAGAAGCTTTCTTTGGCTTGGAAAGAAGATATAACCCAAGAGTTTCCG
GGAAAAGGCAAGGCTTTGTTCCAGGTGAATCAGTGGCTAACCTGGTTAGAACTGCTGAAGAAGAAGAAAT
40 CAGAGGAAGAAGCTGACTAAGAAAGCAGGCCAAAGCCTTAAATTTGTGCAAAACACTACTGTTGCTATGATGT
AATGCAATTTGACCTAACCACTGCAGAAAATTCATTCGCTGTAATGTTTTACAAATATTTAAAGCAGAG
CACGTCAAGTTAGGATTTCTCTTCGATAAGGTTTTTTTTTGTAGTGTAAATGCTTAAATCATAGCTTACCATC
AAATATTTTAGGAGTATCTTTAATGTTTAGATAGTATATTAGCAGCATGCAATAATTACATCATAAGTTT
TCAAGCAGAGGCAGTCTATTGCAAGGACCTTCTTTGCTGCCAGTTATCATAGGCTGTTTTAAGTTAGAAA
45 ACTGAATAGCAACACTGAATACTGTAGAAATGCAC'TTTGCTCAGTAATACTTGAGTTGTTGCAATATTTG
ATTATCCATTTGGTTGTTACAGAAAAATTC'TTAACTGTAATTGATGGTTGTTGCCGTAATAGTATATTGC
CTGTATTTCTACCTCTAGTAATGGGCTTTATGTGCTAGATTTTAAATATCCTTGAGCCTGGGCAAGTGCAC
AAGTCTTTTTTAAAGAAACATGGTTTACTTGCACAAAAC'TGATCAGTTTGGAGAGATCGTTAATGCCCTT
GAAGTGGTTTTTGTGGGTGTGAAACAAATGGTGAGAAATTTGAATTTGGTCCCTCCTATTATAGTATTGAAA
50 TTAAGTCTACTTTAATTTATCAAGTCATGTTTCATGCCCTGATTTTATATACTTGATCTATCAATAAACAT
TGTGATAC'TTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG

SEO ID NO:82

55 >gi|4503539|ref|NP_001409.1| eukaryotic translation initiation factor 4
gamma, 2; Eukaryotic translation initiation factor 4G-like 1; p97 [Homo
sapiens]
MESAIAEGGASRFSASSGGGSGRGAPQHYPKTAGNSEFLGKTPGONAQKWIPARSTRRDDNSAANNSSANE
KERHDAIFRKVRGILNKLTPKFKDKLLLELLNVGVESKLILKGVILLIVDKALEEPKYSSSLYAQLCLRLA
EDAPNEDGPAAEQGQPGQKQSTTFRLLLCLEKQDEFENRTRNVDDVYDKRENPLPEEEEEQRAIAKIMKLG
60 IKFIGELGKLDLILHESILHKCIKLTLEKKRVQLKDMGEDLECLCQIMRTVGPLRDLHERAKSLMDQYFAR
MCSLMLSKELPARIRFLQDQTVELREHHWVPRKAFLDNGPKPTINQIRODAVKDLGVFIAPAPAOGRMSDF

FLEGPFMPPRMKMDRDPGLCLADMFGQMPGSGIGTGPGVIQDRFSPTMGRHRSNQLFNHGGHIMPPTQS
QFGEMGGKFMKSQGLSOLYHNQSQGLLSQLQGQSKDMPPRFSSKKGQLNADEISLRPAQSFLMNKNQVPKL
QPQITMIPPSAQPPRTQTPPLGQTPQLGLKTNPLIQEKPAKTSKKPPPSKEELLKLTETVVTEYLNLSGN
5 ANEAVNGVREMRAPKHFLPEMLSKVIIISLDRSDEDKEKASSLISLLKQEGIIATSDNFMQAFNLVLDQCP
KLEVDIPLVKSYLEAQAARAIISELVSISELAQPLESGTHFPLFLLCLQQLAKLQDREWLTELFQQSKVN
MQKMLPEIDQNKDRMLEILEGKGLSFLFPLLKLEKELLKQIKLDPSPQTIYKWKDNISPKLHVDKGFVN
ILMTSFLQYISSEVNPPSDETDSSSAPSKEQLEQEKQLLSFKPVMQKFLHDHVDLQVSALYALQVHCYN
SNFPKGMLLRFFVHFYDMEIIEEEAFLAWKEDITQEFPGKGKALFQVNQWLWLETAEEEESEEEAD

10 SEQ ID NO:83

Filamin A

>GL3_11C_PCR_G3F1 Direction: sense

ACATCCAGGACACGGTGATGGCAGCACACCATTACCTACATTCCCCTCTGCCCCGGGGCCTACACCGTCACCAT
CAAGTACGGCA

15 SEQ ID NO:84

>GL3_15A_2_M13R Direction: anti-sense

20 TGCCGTACTTGATGGTGACGGTGATAGGCCCGGGGAGAGGGGAATGTAGGTAATGGTGTGCGTGCCATCACCGT
GGTCCTGGATGTTTCGCAGAAACCAGCACACTGGAGCCACCGCTGCACCGGGCTTGGGGGCTCCTCCTGTACAGC
TCGTCCATGCCGAGAGAAGGGCGAATTCTTGACAGATATNCATCACACTGGCGGGCGCTCGAGC

SEQ ID NO:85

>GL3_32C_PCR_G3F1 Direction: sense

25 ACATCAGGACCACGGATGATGGCAGCACACCATACTACATCCCCTCTGCCCCGGGGCCTACACCGTCACCATC
AAGTACGGCACCGNAGAACCAGCACAGTGGTNAGATAGATAAAGCGGCCGCTCGACTAGTCTGAGGTCTGAATA
CTCAACTGAACTGTACGTAAAAACAA
AA
AAAAA

30 SEQ ID NO:86

>gi|4503744|ref|NM_001456.1| Homo sapiens filamin A, alpha (actin binding
protein 280) (FLNA), mRNA

35 GCGATCCGGGCGCCACCCCGCGGTCATCGGTACCGGTCGCTCTCAGGAACAGCAGCGCAACCTCTGCTC
CCTGCCTCGCCTCCCGCGCGCTAGGTGCTGCGACTTTAATTAAAGGGCCGTCCTCGCCGAGGCTGC
AGCACCGCCCCCGGCTTCTCGCGCCTCAAATGAGTAGCTCCCACTCTCGGGCGGGCCAGAGCGCAGC
AGGCGCGGCTCCGGGCGGCGGCTCGACACGCGGGACGCCGAGATGCCGGCCACCGAGAAGGACCTGGCG
GAGGACGCGCCGTGGAAGAAGATCCAGCAGAACACTTTTCACGCGCTGGTGCAACGAGCACCTGAAGTGC
TGAGCAAGCGCATCGCCAACCTGCAGACGGACCTGAGCGACGGGCTGCGGCTTATCGCGCTGTTGAGGT
40 GCTCAGCCAGAAGAAGATGCACCGCAAGCACAACCAGCGGCCCACTTTCCGCCAATGCAGCTTGAGAAC
GTGTCCGTGGCGCTCGAGTTCTTGACCGCGAGAGCATCAAAGTGGTGTCCATCGACAGCAAGGCCATCG
TGGACGGGAACCTGGAAGCTGATCTGGGCCCTCATCTGGACCTGATCTGCACTACTCCATCTCCATGCC
CATGTGGGACGAGGAGGAGGATGAGGAGGCCAAGAAGCAGACCCCCAAGCAGAGGCTCCTGGGCTGGATC
CAGAACAAGCTGCCGAGCTGCCCATCACCAACTTCAGCCGGGACTGGCAGAGCGGGCCGGGCCCTGGGCG
45 CCCTGGTGGACAGCTGTGCCCCGGGCTGTGTCTGACTGGGACTCTTGGGACGCCAGCAAGCCCGTTAC
CAATGCGCGAGAGGCCATGCAGCAGGCGGATGACTGGCTGGGCATCCCCAGGTGATACCCCCGAGGAG
ATTGTGGACCCCAACGTGGACGAGCACTCTGTCTATGACCTACCTGTCCAGTTCCCCAAGGCCAAGCTGA
AGCCAGGGGCTCCCTTGCGCCCCAACTGAACCCGAAGAAAGCCCGTGCTACGGGCCAGGCATCGAGCC
CACAGGCAACATGGTGAAGAAGCGGGCAGAGTTCACTGTGGAGACCAGAAGTGCTGGCCAGGGAGAGGTG
CTGGTGTACGTGGAGGACCCCGGCGACACAGGAGGCAAAAGTGACCGCCAATAACGACAAGAACC
50 GCACCTTCTCCGTCTGGTACGTCCCCGAGGTGACGGGGACTCATAAGGTTACTGTGCTCTTTGCTGGCCA
GCACATCGCCAAGAGCCCTTCGAGGTGTACGTGGATAAGTACAGGGTGACGCCAGCAAAAGTGACAGCC
CAAGGTCCCGGCTGGAGCCAGTGGCAACATCGCCAACAAGACCACCTACTTTGAGATCTTTACGGCAG
GAGCTGGACGGGCGAGGTCGAGGTGTGATCCAGGACCCCATGGGACAGAAGGGCACGGTAGAGCCTCA
GCTGGAGGCCCGGGGCGACAGCACATACCGCTGCAGCTACCAGCCACCATGGAGGGCGTCCACACCGTG
55 CACGTACGTTTGCCGGCGGTGCCCATCCCTCGCAGCCCTACACTGTCACTGTGTTGGCCAAGCCTGTAACC
CGAGTGCTGCCGGGCGGTTGGCCGGGGCCTCCAGCCCAAGGGTGTGCGGGTGAAGGAGACAGCTGACTT
CAAGGTGTACACAAAGGGCGCTGGCAGTGGGGAGCTGAAGGTACCCGTGAAGGGCCCCAAGGGAGAGGAG
CGCTGAAGCAGAAGGACCTGGGGAGTGGCGGTGATGGCTTCGAGTATTACCCCATGGTCCCTGGAACCT
ATATCGTCAACATCACGTGGGGTGGTCAGAACATCGGGCGCAGTCCCTTCGAAGTGAAGGTGGGCACCGA
60 GTGTGGCAATCAGAAGGTACGGGCTGGGGCCTGGGCTGGAGGGCGGCTCGTTGGCAAGTCAGCAGAC

TTTGTGGTGGAGGCTATCGGGGACGACGTGGGCACGCTGGGCTTCTCGGTGGAAGGGUCATCGCAGGCTA
AGATCGAATGTGACGACAAGGGCGACGGCTCCTGTGATGTGCGCTACTGGCCGAGGAGGCTGGCGAGTA
TGCCGTTTACGTGCTGTGCAACAGCGAAGACATCCGCTCAGCCCCCTTCATGGCTGACATCCGTGACGCG
5 CCCCAGGACTTCCACCCAGACAGGGTGAAGGCACGTGGGCCTGGATTGGAGAAGACAGGTGTGGCCGTCA
ACAAGCCAGCAGAGTTTACAGTGGATGCCAAGCACGGTGGCAAGGCCCCACTTCGGGTCCAAGTCCAGGA
CAATGAAGGCTGCCCTGTGGAGGCGTTGGTCAAGGACAACGGCAATGGCACTTACAGCTGCTCCTACGTG
CCAGGAAGCCGGTGAAGCACACAGCCATGGTGTCTGGGGAGGCGTCAGCATCCCCAACAGCCCCCTTCA
GGGTGAATGTGGGAGCTGGCAGCCACCCCAACAAGGTCAAAGTATACGGCCCCGGAGTAGCCAAGACAGG
10 GCTCAAGGCCCCACGAGCCACCTACTTCACTGTGGACTGCGCCGAGGCTGGCCAGGGGACGTCAGCATC
GGCATCAAGTGTGCCCTGGAGTGGTAGGCCCCGCCAAGCTGACATCGACTTCGACATCATCCGCAATG
ACAATGACACCTTACGGTCAAGTACACGCCCCGGGGGGCTGGCAGCTACACCATTATGGTCTCTTTGC
TGACCAGGCCACGCCACCAGCCCCATCCGAGTCAAGGTGGAGCCCTCTCATGACGCCAGTAAGGTGAAG
GCCGAGGGCCCTGGCCTCAGTCGCACTGGTGTGAGCTTGGCAAGCCCCACCCACTTCACAGTAAATGCCA
AAGCTGCTGGCAAAGGCAAGCTGGACGTCCAGTTCTCAGGACTACCAAGGGGGATGCAGTGCGAGATGT
15 GGACATCATCGACCACCATGACAACACCTACACAGTCAAGTACACGCCTGTCCAGCAGGGTCCAGTAGGC
GTCAATGTCACTTATGGAGGGGATCCCATCCCTAAGAGCCCTTTCTCAGTGGCAGTATCTCCAAGCCTGG
ACCTCAGCAAGATCAAGGTGTCTGGCCTGGGAGAGAAGGTGGACGTTGGCAAAGACCAGGAGTTCACAGT
CAAATCAAAGGGTGTGGTGTCAAGGCAAAGTGGCATCCAAGATTGTGGGCCCCCTCGGGTGCAGCGGTG
CCCTGCAAGGTGGAGCCAGGCCCTGGGGGCTGACAAAGTGTGGTGGCTTCTGCCCCGTGAGGAAGGGC
20 CCTATGAGGTGGAGGTGACCTATGACGGCGTGGCCGTGGCAGCCCCCTTCTCTGGAAGCTGTGGC
CCCCACCAAGCCTAGCAAGGTGAAGGCGTTTGGGGCCGGGGCTGCAGGGAGGCAGTGCAGGGCTCCCCGCC
CGCTTACCATCGACACCAAGGGCGCCGGCACAGGTGGCCTGGGCCTGACGGTGGAGGGCCCCCTGTGAGG
CGCAGCTCGAGTGCTTGGACAATGGGGATGGCACATGTTCCGTGTCTACGTGCCCCACCGAGCCCCGGGA
GTACAACATCAACATCCTCTTCGCTGACACCCACATCCCTGGCTCCCCATTCAAGGCCCCAGTGGTTCCC
25 TGCTTTGACGCATCCAAGTCAAGTGTCTCAGGCCCCGGGCTGGAGCGGGCCACCGCTGGGGAGGTGGGCC
AATTCCAAGTGGACTGCTCGAGCGCGGGCAGCGGAGCTGACCATTGAGATCTGCTCGGAGGGCGGGCT
TCCGGCCGAGGTGTACATCCAGGACCACGGTGTATGGCAGCACACCATTAACCTACATTCCCCCTGCCCC
GGGGCCTACACCGTCACCATCAAGTACGGCGGCCAGCCGTGCCCAACTTCCCCAGCAAGCTGCAGGTGG
AACCTGCGGTGGACACTTCCGGTGTCCAGTGCTATGGGCCTGGTATTGAGGGCCAGGGTGTCTTCCGTGA
30 GGCCACCACTGAGTTCAAGTGTGGACGCCCCGGGCTCTGACACAGACCGGAGGGCCGCACGTCAAGGCCGT
GTGGCCAACCCCTCAGGCAACCTGACGAGACCTACGTTACAGGACCGTGGCGATGGCATGTACAAAGTGG
AGTACACGCCTTACGAGGAGGGACTGCACTCCGTGGACGTGACCTATGACGGCAGTCCCGTGCCAGCAG
CCCCCTCCAGGTGCCCGTGACCGAGGGCTGCGACCCCTCCCGGTGCGTGTCCACGGGGCCAGGCATCCAA
AGTGGCACCACCAACAAGCCCAACAAGTTCACTGTGGAGACCAGGGGAGCTGGCACGGGCGGCTGGGCC
35 TGGCTGTAGAGGGCCCCCTCCGAGGCCAAGATGTCTGCATGGATAACAAGGACGGCAGCTGCTCGGTGCA
GTACATCCCTTATGAGGCTGGCACCTACAGCCTCAACGTACCTATGGTGGCCATCAAGTGCAGGCAGT
CCTTTCAAGGTCCCTGTGCATGATGTGACAGATGCGTCCAAGGTCAAGTGTCTTGGGCCCCGGCTGAGCC
CAGGCATGGTTCTGTGCCAACCTCCCTCAGTCTTCCAGGTGGACACAAGCAAGGCTGGTGTGGCCCCATT
GCAGGTCAAAGTGAAGGGCCCCAAGGCCCTGGTGGAGCCAGTGGACGTGGTAGACAACGCTGATGGCACC
40 CAGACCGTCAATTATGTGCCCAGCCGAGAAGGGCCCTACAGCATCTCAGTACTGTATGGAGATGAAGAGG
TACCCCGGAGCCCCCTTCAAGGTCAAGGTGCTGCCTACTCATGATGCCAGCAAGGTGAAGGCCAGTGGCCC
CGGGCTCAACACCACTGGCGTGCCTGCCAGCCTGCCCGTGGAGTTACCATCGATGCAAGGACGCCGGG
GAGGGCCTGTGGCTGTCCAGATCAGGATCCCCAAGGCAAGCCGAAGAAGACACATCCAAGACAACC
ATGACGGCACGTATACAGTGGCCTACGTGCCAGAGTGCAGGTGCGTACACCATCCTCATCAAGTACGG
45 TGGTGACGAGATCCCCTTCTCCCCGTACCGCGTGCCTGCGGTGCCACCGGGGACGCCAAGTGCAC
GTCACAGTGTCAATCGGAGGTACGGGCTAGGTGTGGCATCGGCCCCACCATTCAGATTGGGGAGGAGA
CGGTGATCACTGTGGACACTAAGGCGGCAGGCAAAGGCAAAGTGACGTGCACCGTGTGCACGCTGATGG
CTCAGAGGTGGATGTGGACGTGGTGGAGAATGAGGACGGCACTTTCGACATCTTCTACACGGCCCCCAG
CCGGGCAAATACGTCACTGTGTGCGCTTTGGTGGCGAGCAGTGCCCAACAGCCCCCTTCCAAGTGACGG
50 CTCTGGCTGGGGACAGCCCTCGGTGCAGCCCCCTCTACGGTCTCAGCAGCTGGCCCCACAGTACACCTA
CGCCAGGGCGGCAGCAGACTTGGGCCCCGAGAGGGCCCCCTGGTGGGTGTCAATGGGCTGGATGTGACC
AGCCTGAGGGCCCTTTGACCTTGTCTACCTCCCCCTACCAAGAAGGGCGAGATCACAGGGGAGGTTCCGA
TGCCCTCAGGCAAGGTGGCGCAGCCACCATCACTGACAACAAGACGGCACCGTGACCGTGCAGTATGC
ACCCAGCGAGGCTGGCCTGCACGAGATGGACATCCGCTATGACAACATGCACATCCCAGGAAGCCCCCTTG
55 CAGTTCTATGTGATTACGTCAACTGTGGCCATGTCACTGCCTATGGGCCTGGCCTACCCATGGAGTAG
TGAACAAGCCTGCCACCTTACCCTCAACACCAAGGATGCAGGAGAGGGGGGCTGTCTCTGGCCATTGA
GGGCCCGTCCAAGCAGAAATCAGCTGCACTGACAACAGGATGGGACATGCAGCGTGTCTACCTGCCT
GTGCTGCCGGGGGACTACAGCATTCTAGTCAAGTACAATGAACAGCACGTCCAGGCAGCCCCCTTCACTG
CTCGGGTCACAGGTGACGACTCCATGCGTATGTCCACCTAAAGGTCCGGCTCTGCTGCCGACATCCCCAT
60 CAACATCTCAGAGACGGATCTCAGCCTGTGACGGCCACTGTGGTCCCGCCCTCGGGCCGGGAGGAGCCC
TGTTTGCTGAAGCGGCTGCGTAATGGCCACGTGGGGATTTCATTTCGTGCCCAAGGAGACGGGGGAGCACC

TGGTGCATGTGAAGAAAAAGGCCAGCACGTGGCCAGCAGCCCCATCCCGGTGGTATCAGCCAGTCGGA
AATTGGGGATGCCAGTCGTGTTCTGGGTCTCTGGTCAGGGCCTTCACGAAGGCCACACCTTTGAGCCTGCA
GAGTTTATCATTGATACCCGCGATGCAGGCTATGGTGGGCTCAGCCTGTCCATTGAGGGCCCCAGCAAGG
TGGACATCAACACAGAGGACCTGGAGACCGGACGTGCAGGGTCACCTACTGCCCCACAGAGCCAGGCAA
5 CTACATCATCAACATCAAGTTTGGCCAGCAGTCGCTGGCAGCCCCCTTCTCTGTGAAGGTGACAGGC
GAGGGCCGGGTGAAAGAGAGCATCACCCGAGGCGTCGGGCTCCTTCAGTGGCCAACGTTGGTAGTCATT
GTGACCTCAGCCTGAAAATCCCTGAAATTAGCATCCAGGATATGACAGCCCAGGTGACGACCCATCGGG
CAAGACCCATGAGGCCGAGATCGTGAAGGGGAGAACCACACCTACTGCATCCGCTTTGTTCCCGCTGAG
10 ATGGGCACACACACAGTCAGCGTCAAGTACAAGGGCCAGCAGTCGCTGGGAGCCCCCTTCCAGTTCACCG
TGGGGCCCCCTAGGGGAAGGGGAGCCCAAGGTCCGAGCTGGGGGCCCTGGCCTGGAGAGAGCTGAAGC
TGGAGTGCCAGCCGAATTCAGTATCTGGACCCGGGAAGCTGGTGCTGGAGGCCCTGGCCATTGCTGTGCGAG
GGCCCCAGCAAGGCTGAGATCTCTTTTGGAGACCGCAAGGACGGCTCCTGTGGTGTGGCTTATGTGGTCC
AGGAGCCAGGTGACTACGAAGTCTCAGTCAAGTTCAACGAGGAACACATTCCCGACAGCCCCCTTCGTGGT
GCCTGTGGCTTCTCCGTCTGGCGACGCCCCGCGCTCAGTGTTCCTAGCCTTCAGGAGTCAGGGCTAAAG
15 GTCAACACAGCCAGCCTCTTTTGCAGTCAGCTGAACCGGGCCCAAGGGGGCGATCGATGCCAAGGTGACAC
GCCCCCTCAGGAGCCCCTGGAGGAGTGCTATGTACAGAAATTGACCAAGATAAGTATGCTGTGCGCTTCAT
CCCTCGGGAGAATGGCGTTTACCTGATTGACGTCAAGTTCAACGGTACCCACATCCCTGGAAGCCCCCTTC
AAGATCCGAGTTGGGGAGCCTGGGCATGGAGGGGACCCAGGCTTGGTGTCTGCTTACGGAGCAGGTCTGG
AAGGCGGTGTACAGGGAACCCAGCTGAGTTCTGTCGTGAACACGAGCAATGCGGGAGCTGGTGCCTGTG
20 GGTGACCATTGACGGCCCCCTCCAAGGTGAAGATGGATTGCCAGGAGTGCCCTGAGGGCTACCGCGTCACC
TATACCCCCATGGCACCTGGCAGCTACCTCATCTCCATCAAGTACGGCGGCCCTTACCACATTGGGGGCA
GCCCCCTCAAGGCCAAAGTCACAGGCCCCCGTCTCGTCAGCAACCACAGCCTCCACGAGACATCATCAGT
GTTTGTAGACTCTCTGACCAAGGCCACCTGTGCCCCCAGCATGGGGCCCCGGGTCTTGGGCCCTGCTGAC
GCCAGCAAGGTGGTGGCCGAGGCTGAGCAGGCGCTACGTAAGGCGCTACGTAGGCCAGAAAGACGCTTCACAG
25 TAGACTGCAGCAAAGCAGGCAACAACATGCTGTGGTGGGGGTTTCATGGCCCCAAGGACCCCTGCGAGGA
GATCCTGGTGAAGCACGTGGGCAGCCGGCTCTACAGCGTGTCTTACCTGCTCAAGGACAAGGGGGAGTAC
ACACTGGTGGTCAAATGGGGGCACGAGCACATCCAGGCAGCCCCCTACCGCGTTGTGGTGCCTGAGTCT
GGGGCCCCGTGCCAGCCGGCAGCCCCCAAGCCTGCCCCGCTACCCAAGCAGCCCCGCCCTCTTCCCCCTCAA
CCCCGGCCCCAGGCCGCCCTGGCCGCCCCGCTGTCACTGCAGCTGCCCCCTGCCCTGTGCCGTGCTGCGCTC
30 ACCTGCCTCCCCAGCCAGCCGCTGACCTCTCGGCTTTCACTTGGGCAGAGGGAGCCATTTGGTGGCGCTG
CTTGTCTTCTTTGGTTCTGGGAGGGGTGAGGGATGGG

SEQ ID NO:87

>gi|4503745|ref|NP_001447.1| filamin 1 (actin-binding protein-280); filamin
35 A, alpha (actin-binding protein-280); filamin 1; actin-binding protein-280
[Homo sapiens]
MSSSHSRAGQSAAGAAPGGGVDRDAEMPATEKDLAEDAPWKKIQONTFTRWCNEHLKCVSKRIANLQTD
LSDGLRLIALLEVLSQLKMRKHNRPTFRQMQLENVSVALEFLDRESIKLVSIDSKAIVDGNLKLILGL
IWTILIHYSISMPMWDEEEDDEEAKQTPKQRLGLWIQNKLPQLPITNFSRDWQSGRALGALVDSAPGLC
40 PDWDSWDASKPVTNAREAMQQADDWLGPVITPEEIVDPNVDEHSVMTYLSQFPKAKLKPGAPLRPKLN
PKKARAYGPGIEPTGNMVKKRAEFTVETRSAGQGEVLVYVEDPAGHQEEAKVTANNDKNRTFSVWYVPEV
TGHKVTVLVFAHQHIAKSPFEVYVDKSQGDASKVTAQGPGLPSGNIANKTTYFEIFTAGAGTGEVEVVI
QDPMGQKGTVEPQLEARGDSTYRCSYQPTMEGVHTVHVTFAGVPIPRSPYTVTVGQACNPASACRAVGRGL
QPKGVRVKETADFKVYTKGAGSGELKVTVKGPKGEERVQKDLGDGVYGFYYPMVPGTYIVTITWGGQN
45 IGRSPFEVKVGTECGNQKVRWGPGLGGVVGKSADFVVEAIGDDVGTGLGFSVEGPSQAKIECDDKGDGS
CDVRYWPQEEAGEYAVHVLNSEDIRLSPFMADIRDAPQDFHPDRVKARGPGLKGTGAVNKPAEFTVDAK
HGGKAPLRVQVQDNEGCPVEALVKDNGNGTYSCSYVPRKPVKHTAMVSWGGVSIPNSPFRVNVGAGSHPN
KVKVYGPVAKTGLKAHEPTYFTVDCAEAGQGDVSIKICAPGVVGPAAEADIDFDIIRNDNDTFTVKYTP
RGAGSYTIMVLFAADQATPTSPIRVKVEPSHDASKVKAEGPGLSRTGVELGKPTHFTVNAKAAGKGKLDVQ
50 FSGLTKGDAVRDVIDIDHHDNTYTVKYTPVQQGPVGVNVTYGGDPIPKSPFSVAVSPSLDLSKIKVSGLG
EKVDVGKDQEFVTKSKGAGGQKVASKIVGPSGAAPCKVEPGLGADNSVVRFLPREEGPYEVEVTYDGV
PVPGPSFPLEAVAPTGPSKVKAFGPGGLQGGAGSPARFTIDTKGAGTGGGLGLTVEGPCEAQLECLDNGDG
TCSVSYVPTPEGDYNINILFADTHIPGSPFKAHVPCFADASKVKCSGPGLERATAGEVGQFQVDCSSAGS
AELTIEICSEAGLPAEVYIQDHGDGTHITITYIPLCPGAYTVTIKYGGQVPVPNFPKQLQVEPAVDTSGVQC
55 YGPGIEGQGVFREATTETFSVDARALTQTGGPHVKARVANPSGNLTETVYQDRGDMYKVEYTPYEGLHS
VDVTDGSPVPSSPFPVPTGCDPSRVVHVGPIQSGTTNKNPKFTVETRGAGTGGLGLAVEGPSEAKM
SCMDNKGDCSVEYIPYEAGTYSLVNTYGGHQVPGSPFKVPVHDVTDASKVKCSGPGLSPGMVRANLPQS
FQVDTSKAGVAPLQVKVQGPGLVEPVDVVDNADGTQTVNYVPSREGPYSISVLYGDEEVPRSPFKVKVL
PTHDAKVKASGPGLNTTGVPASLPVEFTIDAKDAGEGLLAVQITDPEGKPKKTHIQDNHDGTYTVAYVP
60 DVTGRYTILIKYGGDEIPFSPYRVRAVPTGDASKCTVTVSIGGHGLGAGIGPTIQIGEETVITVDTKAAG
KGKVTCTVCTPDGSEVDVDVVENEDGTFDIFYTAPQPGKYVICVRFGGEHVPNSPFPQVLTALAGDQPSVQP

PLRSQQLAPQYTYAQQGQQTWAPERPLVGVNGLDVTSLRPFDLVIPFTIKKGEITGEVRMPSGKVAQPTI
TDNKDGTVTTRYAPSEAGLHEMDIRYDNMHIPGSPLQFYVDYVNCGHVTA YGGLTHGVVNKPATFTVNT
KDAGEGGLSLAIEGPSKAEISCTDNQDGTCSVSYLPVLPGDYSILVKYNEQHVPGSPFTARVTGDDSMRM
SHLKVGSAADIPINISETDLSLLTATVPPSPGREEPCLLKRLRNHVGISFVPKETGEHLVHVKKNGQHV
5 ASSPIPVVISQSEIGDASRVVSGQLHEGHTFEPAEFIIDTRDAGYGGLSLSIEGPSKVDINTEDLEDG
TCRVTYCPTPEGNYIINIKFADQHVPGSPFSVKVTGEGRVKESITRRRRAPS VANVGSHCDLSLKIP EIS
IQDMTAQVTPSPGKTHEAEIVEGENHTYCIRFVPAEMGTHTVSVKYKGQHVPGSPFQFTVGP LGE GGAHK
VRAGGPGLERAEGVPAEFSIWTR EAGAGGLAIAVEGPSKAEISFEDRKDGSCGVAYV VQEPGDYEVSVK
FNEEHIPDSPFVVPVSPSGDARRRLTVSSLOESGLKVNQPASFAVSLNGAKGAIDAKVHSPSGALEECYV
10 TEIDQDKYAVRFIPRENGVYLIDVKFNGTHIPGSPFKIRVGEPPGHGGDPGLVSA YGAGLEGGVTGNPAEF
VVNTSNAGAGALSVTIDGPSKVKMDCQECPEGYRVTYTPMAPGSYLSISIKYGGPYHIGGSPFKAKVTGPR
LVSNHSLHETSSVFVDSLTKATCAPQHGAPGPGPADASKVVAKGLGLSKAYVGQKSSFTVDCSKAGNNML
LVGVHGPRTPCEEILVKHVGSR LYSVSYLLKDKGEYTLVVKWGHEHIPGSPYRVVVP

15 SEQ ID NO:88

Gα13 protein

>GL2_46_3_M13R Direction: sense

CCGCGAGGGCAGCGCGCGGCAAGATGGCGGACTTCCTGCCGTGCGGGTCCGTGCTGTCCGTGTGCTTCCCCGG
GNCTGCCTGCTGACGAGTGCGGAGG

20

SEQ ID NO:89

>gi|5729847|ref|NM_006572.1| Homo sapiens guanine nucleotide binding
protein (G protein), alpha 13 (GNA13), mRNA

25 TGGGGCCGAGAGGCGGCGAGGCGGCGGCGGCGGCGGCAAGATGGCGGACTTCCTGCCGTGCGGGTCCGT
GCTGTCCGTGTGCTTCCCCGGCTGCTGCTGACGAGTGCGGAGGCCGAGCAGCAACGCAAGTCCAAGGAG
ATCGACAAAATGCCTGTCTCGGGAAAAGACCTATGTGAAGCGGCTGGTGAAGATCCTGCTGCTGGGCGCGG
GCGAGAGCGGCAAGTCCACCTTCCTGAAGCAGATGCGGATCATCCACGGGCAGGACTTCGACCAGCGCGC
GCGCGAGGAGTTCCGCCCCACCATCTACAGCAACGTGATCAAAGGTATGAGGGTGCTGGTTGATGCTCGA
30 GAGAAGCTTCATATTCCTGGGGAGACAACCTCAAACCAACAACATGGAGATAAGATGATGTGCTTTGATA
CCCCGGGCCCCCATGGCAGCCCCAAGGAATGGTGGAACAAGGGTTTTCTTACAATATCTTCCTGCTATAAG
AGCATTATGGGCAGACAGCGGCATACAGAATGCCTATGACCGGCGTCGAGAATTTCAACTGGGTGAATCT
GTAAAATATTTCTGGATAAATTGGATAAATTGGAGAACCAGATTATATTCCATCACAACAAGATATTC
TGCTTGCCAGAAGACCCACCAAAGGCATCCATGAATACGACTTTGAAATAAAAAATGTTCTTTTCAAAT
35 GCTTGATGTAGGTGGTCAGAGATCAGAAAGGAAACGTTGGTTTGAATGTTTCGACAGTGTGACATCAATA
CTTTTCTTGTTCCTCAAGTGAATTTGACCAGGTGCTTATGGAAGATCGACTGACCAATCGCCTTACAG
AGTCTCTGAACATTTTGAACAATCGTCAATAACCGGGTTTTTCAGCAATGTCTCCATAATTCTGTTCTT
AAACAAGACAGACTTGCTTGAGGAGAAGGTGCAAAATGTGAGCATCAAAGACTATTTCTAGAAATTTGAA
GGGATGCCCCACTGCTTAAGAGACGTCCAAAAATTCCTGGTGGAATGTTTCCGGAACAAACGCCGGGACC
40 AGCAACAGAAGCCCTTATACCACCCTTCACTTCACTATCAACACGGAGAACATCCGCCTTGTTCCTCG
CGACGTGAAGGATACTATTCTGCATGACAACCTCAAGCAGCTTATGCTACAGTGATGTACAAAAGACTTG
CTGTTTTAATATCTTTTTGTGTTTTTGATGTTTTCTGTTTTGTTTTGTTTTTAAATAGCAGTTTACAAC
CAGAATTAGAACAATCTTAATCTACGTTTAACTTCTTGAAAATCTTAGTACTTTTTCTGCGGCCTTTGG
TTTGTGGCTGAAAGCTGTTGAGTGACTCATCGCAAGATTGCTGTAATGCAGGCTTTGATCTGTTTCAC
45 CC

SEQ ID NO:90

>gi|5729848|ref|NP_006563.1| guanine nucleotide binding protein (G
protein), alpha 13; guanine nucleotide binding protein, alpha-13 [Homo
sapiens]

50 MADFLPSRSVLSVCFPGCLLTSGEAEQQRKSKEIDKCLSREKTYVKRLVKILLGAGESGKSTFLKQMRI
IHGQDFDQRRAREEFRPTIYSNVIKMRVLVDAREKLHIPWGDNSNQQHGDKMMSFDTRAPMAAQGMVETR
VFLQYLPALRALWADSGIQNAYDRRREFQLGESVKYFLDNLDKLGEPDYIPSQQDILLARRPTKGIHEYD
FEIKNVFPKMLDVGQRSEKRWFECDVSTISILFLVSSSEFDQVLMEDRLTNRLTESLNIFETIVNNRV
FSNVSIILFLNKTDLLEEKVQIVSIKDYFLEFEGDPHCLRDVQKFLVECFRNKRDRDQQQKPLYHHFTTAI
55 NTENIRLVFRDVKDITLHDNLKQLMLQ

SEQ ID NO:91

HOXB2

>GL3_11D_1_M13R Direction: sense

5 TTTGCCCTGACCCGGGCCATTGCCAGAAGACGTCTTCTCGGGGCGCCAGGATTCACCTTTCCTTCCCGACCTCAA
CTTCTTCGCGGCCGACTCCTGTCTCCAGCTATCCGGAGGCCTCTCCCCCTAGCCTACAGGGTTCTCTCGACAGCCC
TGGTCCCTTTTCCGATGGACAAAGAGCTTGGATTTTTTCACCAAGTACCGCTCTGTGCCCAATTCCGGAGCCTGC
NAGGTTTTACCCCCCCTTTAAAAAGCCCCCTTTGGTTATATACCCCTTACCCTTACCCCCCGTGGGTTACCCC
10 TTNTATACGGAACCCCCCTGGTCGGCCTTCCCCTTTATGGGGCCCCGGTCTTAAACCTGGGGAAACACAAN
AATTCCGGAAGGGACCCCCCATTTTCGATTTACCCCGCAAAAGCCCCCNCTTTCCCAAANGGATACCGGCCA
AGTTNTAAAGGAAACCTTTTAAANATTTGGATTGAGGATTATTTATTATAGGGGCCCTATTAAAAACAAACAACAA
CACTTCTTTTTATCCACAAAGTGGNGGGATNTAAATGATTTATCAAANCCCGTTTGTGGTAACAAAACAATTTT
TTTACAAAACGNCCCGGGAGTATATTAGTTATAAAAAAATATTCCTCCCGAAAAACATTATTTCTCCCCCCCCC
CTTATATATTTTAAACACCCCATATATTCAACNGCTATTATTATTGCCACCATATTTTACCCCATAAATA

15 SEQ ID NO:92

>GL3_21C_1_M13F Direction: N/A

TTTGCCCTGACCCGGGCCATTGCCAGAAGACGTCTTCTCGGGGCGCCAGGATTCACCTTTCCTTCCCGACCTCAA
CTTCTTCGCGGCCGACTCCTGTCTCCAGCTATCCGGAGGCCTCTCCCCCTTAGCCTACAGGGTTCTCTCGACAGC
20 CCGTCCCGGTTTTTTATTTCCCGGAAGGGGGAAGGGAAGGCTGGGGATTGTGTTTCGACGCAGTACGCTCTGT
GCCATCGACCTGCAGTTTCCCTAACCTGTTTCTCTCCCGGTCCTTTCGACCCCCGCGCTCCTTGCGCGTCTAC
TGGA

SEQ ID NO:93

>GL3_11D_1_M13F Direction: anti-sense

25 AAAANGTATAGACTCACTATAGTGGCGAATTGGGNCCATCTAGATGCATGCTCGAGCGGCCGCCAGGTGTGATGG
ATATCTGCAGAATTGCTGCCGCTTACGACAGTCAGTGAGTATCAGACCTCAGAGCTAGATCGAGCGGCCGCTTTA
TCTATCTAACCAC'TGTGCTGGGTTTCGTGCGGNCCCCGCTAGAGTTTAAATGTATTCTGAGATTTCACTGGACAGG
AGTCTACCAAACGGAATTTTTCTGTGTGAATTTTAAAGATAACCGAGTGCCCAATATTTTAGAACGAAGAAGAA
30 AGGGAGTGGATTAAACGCTAAGTTTCAGTAATACCNCTGAAGTTTTTAGCAAAAGCGACATAAGTTCTATGGCGACT
GAGGGGTGGGAGAGGCTCGACGNNTTNAACCAAGTAGACGGGCCAAGGAAGGCGCGNGGGGTCCGGAAACANGG
GACCCGGGGCAGGGGGAGGGGAAAAAACAGGGTTACAGCGNGAAAAACCTGGCCAAGGGACTACCCGGAACGTAT
GAGGGCCAAACAAAAGAAAGGCCGGATTAAACCTATGGGGGATTGGAAAAAAAATCCCAAAGGGCTTCTAT
ATCCCCCTCCGGGGGAAAAAACACAGGGGGGAAACCCCCGGGAGGCCCTATGGATCCAGGACACGAAAGCGAA
35 AACCCCCCTGGGTAAGAGAGCCCTAAGGGGGGA

SEQ ID NO:94

>GL3_27D_2_M13R Direction: anti-sense

CCCCCTAGAGTTTAATTATTCCTGAGATTTCACTGGAAGGAGTCTACCAAACGGAATTTTTCTGTGTGAATTTT
AAAAGATAACCGAGTGCCCAATATTTTAGAAGAAGAAGAAAGGGAGTGGATTAAACGCTAATTCAGTAATACCTG
40 AATTTTAGCAAAACACATAAGTCTATGCGACTGAGGGTGGGAGAGGCTCGATTTTTCCAGTAGACGGCCAAGGAG
CGCGGGGTGCGAAAGGACCGGAGGAGGAAGACAGGTTAGGGAAACTGCAGGTCGATGGCACAGAGCGTACTGGTG
AAAAATCCAGCTCTTCTCGGAAAAAGGGACCGGGCTGTGAGAGAAACCTGTAGGCTAGGAGGAGGACCTCCG
GATAGCTGGAGACAGGAGTCCGCCGCGAAGAAGTTGAGGTCCGGAAGGAAAGGTGAATCCTGGCGCCCCGAGAAG
45 ACGTCTTCTGGCAATGGCCCCGGGCTCCAGCCCCACGGCCCCGCGCAGCGCGCAAGCCGGGACTTCGAACGCACNC
NTGCAGCCCTCATAAGCGAACGGCATAAAAGGCCCGGGGGTCCAGCGCATTAAAGGGCCCCCGGACCATCCTCC
GGGCGAGGTGAACAAGCGGGTTACCCAAGGCCGGCGAGAGGTCCCGAGGCCATTNGACTAGGGNT

SEQ ID NO:95

>GL3_45B_3_M13F Direction: N/A

50 TTTGCCCTGACCCGGGCCGGGGCTGTAGACTAGAGAGGGAGAGAAGAAAAACCCAAAAGCAATGGCCTGGGACTGC
CGGAGGCTGGTGGCGCGGGGCGCGCAGGCTGCGCACGGGTTTACACCAACACGCAGCTGCTGGAACCTGGAGAAG
GAATTCCACTTTAATAAGTACCTGTGCGGCCACGCCGCTCGAGATCGCGGCCCTTGCTGGACCTCACCAGAAAGG
CAGGTCAAAGTCTGGTTTTCAGAACCGGCGCATGAAGCACAAGCGGCAGACGCAGCACCGGAGCCGGATGGG
55 GAGCCTGCCTGCCCGGGAGCCCTGGAGGACATCTGCGACCCTGCCGAGGAACCTGCGGGCCAGCCGGGCGGCC
CTCCGCTCGCGGGCGCGTGGGAAGCCTGCTGTACCCGCCGGAGGTGGTTGCCGGGGGCC

SEQ ID NO:96

>GL3_27D_1_M13R Direction: anti-sense

60 CCCCCCTAGAGTTTAATTATTCCTGAGATTTCACTGGAAGGAGTCTACCAAACGGAATTTTTCTGTGTGAATTTT
AAAAGATAACCGAGTGCCCAATATTTTAGAAGAAGAAGAAAGGGAGTGGATTAAACGCTAATTCAGTAATACCTG

AATTTTAGCAAAACACATAAGTCTATGCGACTGAGGGTGGGAGAGGCTCGATTTTTCCAGTAGACGGCCAAGGAG
CGTGGGGGTGCAAAAGGACCGGGAGGAGGAAACAGGTTAGGGAAACTGCAGGTCGATGGCACAGAGCGTACTGGTG
AAAAAATCCAGCTCTTCTCGGAAAAAGGGACCGGGCTGTGAGAGAACCTGTAGGCTAGGGGAGAGGCCTCCG
GATAGCTGGAGACAGGAGTCGGCCGCGAAGAAGTTGAGGTCGGGAAGGAAAGGTAATTCCTTCTCCAGTTCCAGC
AGCTGCGTGTGGTGTAAAGCCGTG

SEQ ID NO:97

>GL3_45B_1_M13R Direction: sense

TTTGCCCTGACCCGGGCGGGGCTGTAGACTAGAGAGGGAGAGAAGAAAACCCAAAAGCAATGGCCTGGGACTGC
CGGAGGCTGGTGGCGGGCGGGCGCGCAGGCTGCGCAGGTTACACCAACACGCAGCTGCTGGAACGGAGAAGGA
ATTCCACTTTAATAAGTACCTGTGCCGGCCACGCCGCTCGAGATCGCGGCTTGCTGGACCTACCGAAAGGCA
GGTCAAAGTCTGGTTTTCAGAACCGGCGCATGAAGCACAAGCGGCAGACGACGACCGAGAGCCGCGGATGGGGA
GCCTTGCCCTTGCCGGGAGCCCTTGAGGACATCTGCGACCCTGCCGAGGAACCCGCGGGCCAGCCCGGGCGGGC
CCTTCCGCCTCGCGGGCGGGCG

SEQ ID NO:98

>GL3_11D_3_M13R Direction: anti-sense

CCCCCTAGAGATTCTAATGTACTTCTGAGATTATCACTGGAAGGAGATCTACCAAACGGAATTTCTTCCGTGT
GAATTTTAACAGAGATAACCGAGTGCCCATATTTTAGAAGAAGAAGAAAGGGAGTGGATTAAACGCTAATTCAGT
AATACCTGAATTTTAGCAAAGACACATAAGTCTATGCGACTGAGGGTGGGAGAGGCTCGATTTTTCCAGTAGACG
GCCAAGGAGCGCGGGGTCGAAAGGACCGGGAGGAGGAAACAGGTTAGGGAAACTGCAGGTCGACTGGCACAGAG
CGTACTGGTGAAAAAATCCAGCTCTTCTCGGAAAAAGGGACCGGGCTGTGAGAGAACCCTGGTAGGCTAGGGG
AGAGGCCTCCGATAGCTGGAGACAGGAGTCGGCCGCGAANGAAGTTGAGGTCGGGAAGGAAAGGTGAATCC

SEQ ID NO:99

>GL3_11D_4_M13F Direction: sense

AAATTGGAAATAGACTCACTATAGGGCGAATTGGGCCCTCTAGAATGACATGCTCGAGCGGCCAGNCCAAGGTGT
GATGGATATCTGCAGAATTCTGCCCTTACGACAGTCAGTGAGTATCAGACCTCAGAGCTAGATCGTAGCGGCCGC
TTTATCTATCTAACCCTGTGCTGGGTNTCTGCGGCCCCCGCTAGAGTTTAATTATTCCTGAGATTTCACTGGCC
AGGGAGTCTACCAACAGGGAATTTTCTGTGTGAATTTTAAAAGATAACCGAGTGCCCAATATTTTAGAAGAAGA
AGAAAGGAGTGGATTAAACGCTAATTCAGTAATACCTGAATTTTAGCAAAACACATAAGTCTATGCGACTGAGG
GTGGGAGAGGCTCGATTTTCCAGTAGACGGCCAAGGAGCGCGGGGNCGAAAGGACCGGGAGGAGGAAACAGGT
TAGGGAAACTGCAGTTCGATGGCACAGAGCGTACTGGTGAAAAAATCCAGCTCTTCTCGGAAAAAGGGACCGGG
GTGTAAGAGAGAACCCTGGNAAGGCCAGCGGGA

SEQ ID NO:100

>GL3_41D_1_M13R Direction: anti-sense

CCCCCTAGAGTTTAATTATTCCTGAGATTTCACTGGAAGGAGTCTACCAACGGAATTTTTCTGTGTGAATTTT
AAGAAGATAACCGAGTGCCCAATATTTTAGAAGAAGAAGAAAGGGAGTGGATTAAACGCTAATTCAGTAATACCT
GAATTTTAGCAAAACACATAACGTCTATGCGACTGAGGGTGGGAGAGGCTCGATTTTGTCCAGTACGACGGCCAA
GGAGCGCGCGGGGTCGAAAGGACCGGGAGGAGGAAACAGGTTAGGGAAACTGCAGGTCGATGGCACAGAGCGTAC
TGGTGAAAAATCCAGCTCTACCTCGGAAAAAGGGACCGGGACCNATCGAGAGAACCCTGTAGGCTAGGGGAGA
GGCCATCCGGATAGACTGGAGACAGGAGTCGGAACGCGAAGAAGTACGAGGTCGNGGAAGG

SEQ ID NO:101

>GL3_41D_3_M13R Direction: anti-sense

CCCCCTAGAGTTTAATTATTCCTGAGATTTCACTGGAAGGAGTCTACCAACGGAATTTTTCTGTGTGAATTTT
AAAAGATAACCGAGTGCCCAATATTTTAGAAGAAGAAGAAAGGGAGTGGATTAAACGCTAATTCAGTAATACCTG
AATTTTAGCAAAACACATAAGTCTATGCGACTGAGGGTGGGAGGGGCTCGATTTTCCAGTAGACGGCCAAGGAG
CGCGGGGTGCAAAAGGACCGGGAGGAGGAAACAGGTNAGGGAAACTGCAGGTCGATGGCACAGAGCGTACTGGTG
AAAAAATCCAGCTCTTCTCGGAAAAAGGGACCGGGCTGTGAGAGAACCTGTAGGCTAGGGGAGAGGCCTCCG
GATAGCTGGAGACAGGAGTCGGCCGCGAAGAAGTNGAGGTGCGGAAGGAAAGCGTGAAAGTCCCTGGCACAGCCA
CCGAAGGCAAGAACGGATCNTTTTCATNGGACAAAAATGGGACCCAGGAGCATCCCAAGCACCCGGANCGGGGCG
CCCCGGACGCANGCAGNAGCGCAAGACAGGGNGAATCAGGAAAGGCAGNACNTGACGACCCCATTCAAAAAG
GAGAGAACAANGGNATAAAAAGGGACCAGGAGAAGANCAAG

SEQ ID NO:102

>GL3_41D_4_M13R Direction: anti-sense

CCCCCTAGAGTTTAATTATTCCTGAGATTTCACTGGAAGGAGTCTACCAACGGAATTTTTCTGTGTGAATTTT
AAAAGATAACCGAGTGCCCAATATTTTAGAAGAAGAAGAAAGGGAGTGGATTAAACGCTAATTCAGTAATACCTG

5

>GL3_45B_1_M13F Direction: anti-sense

10

>GL3_45B_3_M13R Direction: anti-sense

20

>GL3_45B_4_M13R Direction: anti-sense

30

>GL3 27D 3 M13F Direction: N/A

40

>GL3 27D 3 M13R Direction: anti-sense

50

>GL3 27D 4 M13F Direction: N/A

60

5

>GL3_27D_4_M13R Direction: anti-sense

10

>GL3_21C_1_M13R Direction: anti-sense

20

>GL3_41D_3_M13F Direction: N/A

30

>GL3 41D 4 M13F Direction: N/A

40

>GL3 41D PCR G3F1 Direction: N/A

50

SEQ ID NO:114

>GL3_41E_PCR_G3F1 Direction: sense

TTTGCCCTGACCCGGGCGGGGCTGTAGACTAGAGAGGGAGAGAAGAAAACCCAAAAGCAATGGCCTGGGACTGC
CGGAGGCTGGTGGCGGGCGGGGCGCGCAGGCTGCGCACGGCTTACACCAACACGCAGCTCTGGAAGTGGAGAAGGA
ATTCCACTTTAATAAGTACCTGTGCCGGCCACGCCGCTCGAGATCGCGGCCTTGCTGGACCTCACCAGAAAGGCA
GGTCAAAGTCTGGTTTCAGAACCGGCGCATGAAGCACAAGCGGCAGACGCAGCACCAGAGAGCCGCCGGATGGGGA
GCCTGCCTGCCCGGGAGCCCTGGAGGACATCTGCGACCCCTGCCGAGGAACCCGCGGCCAGCCCGGGCGGGCCCTC
CGCCTTCGCGGGCGGCGTGGGAAGCCTGCTGTACCCGCGCGAGGTGGTGGCGGGGGCCTTAAGCGCGGACCCCCG
GCCTTTAGCCGTTTCGCTTAGAGGGCGCAGGCGCGTTCGAGTCCCGGCTGCGCGCTGCGCGGGGCGGCGGGCTGGA
GCCCGGGCCATTGCCAGAAGACGTCTT

SEQ ID NO:115

>GL3_45B_4_M13F Direction: N/A

TTTGCCCTGACCCGGGCGGGGCTGTAGACTAGAGAGGGAGAGAAGAAAACCCAAAAGCAATGGCCTGGGACTGC
CGGAGGCTGGTGGCGGGCGGGGCGCGCAGGCTGCGCACGGCTTACACCAACACGCAGCTGCTGGAAGTGGAGAAGG
AATTCCACTTTAATAAGTACCTGTGCCGGCCACGCCGCTCGAGATCGCGGCCTTGCTGGACCTCACCAGAAAGGC
AGGTCAAAGTCTGGTTTCAGAACCGGCGCATGAAGCACAAGCGGCAGACGCAGCACCAGAGAGCCGCCGGATGGGG
AGCCTGCCTGCCCGGGAGCCCTGGAGGACATCTGCGACCCCTTGCCGAGGAACCCGCGGCCAGCCCGGGCGGGCC

SEQ ID NO:116

>GL3_45B_PCR_G3F1 Direction: N/A

TTTGCCCTGACCCGGGCGGGGCTGTAGACTAGAGAGGGAGAGAAGAAAACCCAAAAGCAATGGCCTGGGACTGC
CGGAGGCTGGTGGCGGGCGGGGCGCGCAGGCTGCGCACGGCTTACACCAACACGCAGCTGCTGGAAGTGGAGAAGG
AATTCCACTTTAATAAGTACCTGTGCCGGCCACGCCGCTCGAGATCGCGGCCTTGCTGGACCTCACCAGAAAGGC
AGGTCAAAGTCTGGTTTCAGAACCGGCGCATGAAGCACAAGCGGCAGACGCAGCACCAGAGAGCCGCCGGATGGGG
AGCCTGCCTGCCCGGGAGCCCTGGAGGACATCTGCGACCCCTGCCGAGGAACCCGCGGCCAGCCCGGGCGGGCCCT
CCGCCTCGCGGGCGGCGTGGGAAGCCTGCTGTACCCGCGCGAGGTGGTGGCGGGGGCCTTAAGCGCGGACCCCC
GGCCTTTAGCCGTTTCGCTTAGAGGGCGCAGGCGCGTTCGAGTCCCGGCTGCGCGCTGCGCGGGGCGGCGGGCTGG
AGCCCCGGGCCATTGCCAGAAGACGTCTTCTCGGGGCG

SEQ ID NO:117

>GL3_27D_1_M13F Direction: N/A

TTTGCCCTGACCCGGGCGGGGCTGTAGACTAGAGAGGGAGAGAAGAAAACCCAGAAAGCAATGGCCTGGGACTGC
CGGAGGCTGGTGGCGGGCGGGGCGCGCAGGCTTGC GGCCAACGGTTTTACACCAAAACACGCAGCTGCTTGGAA
CTGGAGAAAGGAAATTAACCTTTTTCTTTCCCGAACCTTCAAACCTTTCTTTTCGGCGGGCCGGACTTCTTTGTC
TCCAGCTTATCCGGAGAGCGCTTCTCCCCATTAGCCCTAACAAAGGGGTTCTCTCNGGACAGGCCCGCGGGGTCC
CCTTTTTTCCGGAGAAAGAGCTTGAATTTTCCAACCAAGGTACGGCCATTCTGTTGGTTGCCCCCAATTTTCGT
AGAACCCTTTGCACAAGGTTTTTTCCCCCTATAAAAACCTGTGGTTTTTTCCCTTCCACCTTTCCCGCGGGGG
TATCCCTTTTTTTCAGGAAGCCGCGCCNACAAAAGGGCCTTCTATTTTGGGCCCCGGTGTACTCTTTAGCGTGTGG
GAAAACAGAATTCGTGCGGAAGGCCCTTCTTCCACAAACCCCCATTCTNAGAGGTCCGGGGCAAGATTTAGAAA
CTTTATTGGACTATGATTTGTGTTGAGGGCCCTTAAACACAAAATTTTCAAGGGGATTATTTTTTA

SEQ ID NO:118

>GL3_27D_2_M13F Direction: N/A

TTTGCCCTGACCCGGGCGGGGCTGTAGACTAGAGAGGGAGAGAAGAAAACCCAAAAGCAATGGCCTGGGCACTG
CCTGGCGCTTTGGGAAGCGTGCTGTACCCGTCGGAGGGTGGATGCCGTGGGGTCTTATGAAGCTGCGGACCCC
CGGTCTTTAGACCGTACGCTTAGAGCGGCGGACAGGCNCGTCGAGGTCCCGGCTGCGAACAGCGCGGAGCCCGGG
CNGCATGAAGCCTCGTGTNCNAATTTNCCAAGAATGGAAGGGTTCTTACCCTTCNGAGGCGCGGCCCAAGNGTA
CTTCTTACCCCTGTGTTTCCCTTTTCCCCGAAACCTCTCGAAGTCTTTCTTTACGGGCGNGGCCAGANACTC
CGTTGTTCTTACCCTANGGATTACANCTGAGAAGACCTTTTNCNACGCCTCNTTGAGGCCTCACTGTAAACNGG
TGGGAGGGTTNGTCTCCGTCTTNCGGGGACCTAAGCGGCCCGCTGGGCGAACCCGGTTTTTTTTAAACCTCA
CCGGGAGAAAGNAGACAAAANGACAAAAGGGTCTTTTGAGGAAGATTCCATCTTTNACNACNCCGTATAGATTAGT
TAACGGGTCTCCCTTGTGCTTCGCCCACATATTTAAGTACCGGCTGTGGGCGAAGCGTTCTTGGCCCNCCAAATTA
CACAGTCCNCTTTGTGATATAACCCNCTTCCACTCTTACCCGNGCGGGGACTCCCACTTTTTTTTTC

SEQ ID NO:119

>GL3_41D_1_M13F Direction: N/A

TTTGCCCTGACCCGGGCGGGGCTGTAGACTAGAGAGGGAGAGAAGAAAACCCAAAAGCAATGGCCTGGGACTGC
CGGAGGCTGGTGGCGGGCGGGGCGCGCAGGCTGCGCNACGGTTACACTCAACTACGCAGCTGCTGGAAGTGGAGA
AGGAATTTCCACTTTAATAAGTACCTTGTGTCCGGGCCACAGCCGGCGGTTCAGAGATCNGCGGGCCTTTGCCTGG

ACCCTTCAGCCCCGAAAAGGGCAAGGTCNAAGAGGCTCATAGGGTGTCCAGAGACNCGGNGGCGGCCAATTGAAGA
GGCAACGAGAGGCCGGCCAAGGAACGGCAGCAACCCGGAAGAAGGCCCGGCCGNGGCATTGGGGGGGAAGCCTT
NGCCCTTGCCCCGGTGGTGAAGGTCCGCTTTGGGAANGGGAGCATATTCCTNGCGGAAACGCCTTGTCNCAGAA
GGGTAGAAAACCCGTGACGGGGCCAAGGGCGCCCCGGGGGCCGGCGGCCCCCCCTTTCCGGACCCATTTCCGGGGCC
5 GGGGNGCCGGNGCAGTTG

SEQ ID NO:120

>gi|4504464|ref|NM_002145.1| Homo sapiens homeo box B2 (HOXB2), mRNA
GAATTCGGGGGGGGTCCCTTCCGATCCTCCCTCCTGACGCCCCCCCCAGCAGCCCCCTCCCCACCAT
10 TGAAAGCCATGAATTTTGAATTTGAGAGGGAGATTGGGTTTATAAACAGCCAGCCGTCGCTCGCCGAGTG
TCTGACTTCCCTTCCCCGCTGTCTTGGAGACATTTCAAACCTTCATCAATCAAGGAGTCGACATTAATTCCT
CCTCCTCCTCCTTTTCGAGCAAACCTTCCCCAGCCTCCAGCCCCGGCGCCTCCACCCTTCAGAGACCCAGGA
GCCAAAAGCGAGCCGAAGATGGGCCTGCTCTGCCGCCGCCACCGCCGCCCACTCCCCGCTGCCCCCCC
15 GGCCCCCGAGTTCCCTTGGATGAAAGAGAAGAAATCCGCCAAGAAACCCAGCCAATCCGCCACGTCTCCT
TCTCCGGCCGCTCCGCCGTTCCGGCCTCCGGGGTCCGATCGCCTGCAGATGGCCTGGGACTGCCGGAGG
CTGGTGGCGGGCGGGGCGCGCAGGCTGCGCACGGCTTACACCAACACGCAGCTGCTGGAACCTGGAGAAGGA
ATCCACTTTAATAAGTACCTGTGCCGGCCACGCCGCTCGAGATCGCGGCCTTGCTGGACCTCACCGAA
AGGCAGGTCAAAGTCTGGTTTTCAGAACCGGCGCATGAAGCACAAGCGGCAGACGCAGCACCGAGAGCCGC
20 CGGATGGGGAGCTGCTTCCCCGGGAGCCCTGGAGGACATCTGCGACCCTGCCGAGGAACCCGCGGCCAG
CCCCGGCGGCCCTCCGCTCGCGGGCGGCGTGGGAAGCCTGCTGTCACCCGCGGAGGTGGTGGCGGGG
GCCTTAAGCGCGGACCCCCGGCCTTTAGCCGTTTCGCTTAGAGGGCGCAGGCGCGTTCGAGTCCCGGCTGCG
CGCTGCGCGGGGCGGGCGGGCTGGAGCCCGGGCCATTGCCAGAAGACGTCTTCTCGGGGCGCCAGGATTC
ACCTTTCCTTCCCGACCTCAACTTCTTCGCGGCCGACTCCTGTCTCCAGCTATCCGGAGGCCTCTCCCCCT
25 AGCCTACAGGGTCTCTCGACAGCCCGGTCCCTTTTTCCGAGGAAGAGCTGGATTTTTTACCAGTACGC
TCTGTGCCATCGACCTGCAGTTTCCCTAACCTGTTTCTCCTCCCGTCCCTTCGACCCCCGCGCTCCTT
GGCCGTCTACTGGAAAAATCGAGCCTCTCCACCCTCAGTCGCATAGACTTATGTGTTTTGCTAAATTC
AGGTATTACTGAATTAGCGTTTAATCCACTTCCTTTCTTCTTCTTCTAAAATATTGGGCACTCGGTTATC
TTTTAAAATTACACAGAAAAATCCGTTTGGTAGACTCCTTCCAATGAAATCTCAGGAATAATTAAACT
30 CTAGGGGGACTTTCTTAAAAATAACTAGAGGGACCTATTTTCTTCTTTTATGTTTTAGACTGTAGATT
ATTTATTAAATCTTTAATAATAGGAAAAGGGGAAAGTATTTATTGTAC

SEQ ID NO:121

>gi|4504465|ref|NP_002136.1| homeo box B2; Homeo box-B2 [Homo sapiens]
MNFEREIGFINSQPSLAELTSFPAVLETFQTSSIKESTLIPPPPPFEQTFPSLQPGASTLQRP RSQK
35 RAEDGPALPPPPPPPLPAAPPAPPEFPWMKEKKS AKKPSQSATSPSPAASAVPASGVGSPADGLGLPEAGG
GGARRLR TAYTNTQLLELEKEFHFNKYLCPRRRVEIAALLDLTERQVKVWFQNRMRMKHRQTQHREPPDG
EPACPGALEDICDPAEPAASPGGPSASRAAWEACCHPPEVVP GALSDPRPLAVRLEGAGASSPGCALR
GAGGLEPGPLPEDVFSGRQDSPFLPDLNFFAADSCLQLSGGLSPSLQGS LDSPVPFSEELDFFTSTLCA
40 IDLQFP

SEQ ID NO:122

HSPA5

>GL2-67-2M13R Direction: N/A

CCGATGAGCCGCTTGGCGTCAAAGACCGTGTCTCGGGGTTGGAGGTGAGCTGGTTCTTGGCGGCATCGCCAATC
45 AGACGTTCCCCTT CAGGAGTGAAGGCGACATAGGACGGCGTGATGCGGTTGCCCTGATCGTTGGCGATGATCTCC
ACGCGGCCGCTCGCGATCTAGAAC

SEQ ID NO:123

>GL2-67-5M13F Direction: N/A

GTTCTAGATCGCGAGCGGCCGCGTGGAGATCATCGCCAACGATCAGGGCAACCGCATCACGCCGTCTATGTGCG
50 CTTCACTCCTGAAGGGGAACGTCTGATTGGCGATGCCGCCAAGAACCAGCTCACCTCCAACCCGAGAACACGGT
CTTTGACGCCAGCGGCTCATCGG

SEQ ID NO:124

>GL2-67-5M13R Direction: N/A

CCGATGAGCCGCTTGGCGTCAAAGACCGTGTCTCGGGGTTGGAGGTGAGCTGGTTCTTGGCGGCATCGCCAATC
55 AGACGTTCCCCTT CAGGAGTGAAGGCGACATAGGACGGCGTGATGCGGTTGCCCTGATCGTTGGCGATGATCTCC
ACGCGGCCGCTCGCGATCTAGAAC

SEQ ID NO:125

>GL2_1_A03_G3F1 Direction: sense

GTTCTAGATCGCGAGCGGCCGCGTGGAGATCATCGCCAACGATCAGTGGCAACCGCATCACGCCGTCTTATGTCTG
CCTTCACTCCTGAAGGGGAACGTCTGATTGGCGATGCCGCCAAGAACCAGCTCACCTCCAAACCCGAGAACACGG
TCTTTGACGCCAAGCGGCTCATCTGG

SEQ ID NO:126

>GL2_1_A05_G3F1 Direction: sense

GTTCTAGATCGCGAGCGGCCGCGTGGAGATCATCGCCAACGATCAGGGGCAACCGCATCACGCCGTCTTATGTCTG
CCTTCACTCCTGAAGGGGAACGTCTGATTGGCGATGCCGCCAAGAACCAGCTCACCTCCAAACCCGAGAACACGG
TCTTTGACGCCAAGCGGCTCATCTGG

SEQ ID NO:127

HSP5>

ATGGAGGAGGACAAGAAGGAGGACGTGGGCACGGTGGTTCGGCATCGACCT
GGGGACCACCTACTCCTGCGTGGCGTGTTCAGAACGGCCGCGTGGAGA
TCATCGCCAACGATCAGGGCAACCGCATCACGCCGTCTTATGTCTGCCTTC
ACTCCTGAAGGGGAACGTCTGATTGGCGATGCCGCCAAGAACCAGCTCAC
CTCCAACCCCGAGAACACGGTCTTTGACGCCAAGCGGCTCATCGGCCGCA
CGTGGAATGACCCGTCTGTGCAGCAGGACATCAAGTTCTTGCCGTTCAAG
GTGGTTGAAAAGAAAATAAACCATAATTCAAGTTGATATTGGAGGTGG
GCAAACAAAGACATTTGCTCCTGAAGAAATTTCTGCCATGGTTCTCTACTA
AAATGAAAGAAACCGCTGAGGCTTATTTGGGAAAGAAGGTTACCCATGCA
GTTGTTACTGTACAGCCTATTTTAATGATGCCCAACGCCAAGCAACCAA
AGACGCTGGAACATATTGCTGGCCTAAATGTTATGAGGATCATCAACGAGC
CTACGGCAGCTGCTATTGCTTATGGCCTGGATAAGAGGGAGGGGGAGAAG
AACATCCTGGTGTGTTGACCTGGGTGGCGGAACCTTCGATGTGTCTCTTCT
CACCATTGACAATGGTGTCTTTCGAAGTTGTGGCCACTAATGGAGATACTC
ATCTGGGTGGAGAAGACTTTGACCAGCGTGTCTATGGAACACTTCATCAAA
CTGTACAAAAAGAACGCGGCAAGATGTCAGGAAAGACAATAGAGCTGT
GCAGAAACTCCGGCGCGAGGTAGAAAAGGCCAAACGGGCCCTGTCTTCTC
AGCATCAAGCAAGAAATTGAAATTGAGTCTTCTATGAAGGAGAAGACTTT
TCTGAGACCCCTGACTCGGGCCAAATTTGAAGAGCTCAACATGGATCTGTT
CCGGTCTACTATGAAGCCCGTCCAGAAAGTGTGGAAGATTCTGATTTGA
AGAAGTCTGATATTGATGAAATTGTTCTTGTGTTGGTGGCTCGACTCGAATT
CCAAAGATTTCAGCAACTGGTTAAAGAGTTCTTCAATGGCAAGGAACCATC
CCGTGGCATAAAACCCAGATGAAGCTGTAGCGTATGGTGTCTGTGTCAGG
CTGGTGTGCTCTCTGGTGTATCAAGATACAGGTGACCTGGTACTGCTTGAT
GTATGTCCCCCTTACACTTGGTATTGAAACTGTGGGAGGTGTCTATGACCAA
ACTGATTCGAAGAACACAGTGGTGCCTACCAAGAAGTCTCAGATCTTTT
CTACAGCTTCTGATAATCAACCAACTGTTACAATCAAGGTCTATGAAGGT
GAAAGACCCCTGACAAAAGACAATCATCTTCTGGGTACATTTGATCTGAC
TGGAATTCTCTCTGCTCCTCGTGGGGTCCACAGATTGAAGTCACCTTTG
AGATAGATGTGAATGGTATTCTTCGAGTGACAGCTGAAGACAAGGGTACA
GGGAACAAAAATAAGATCACAATCACCATGACCAGAATCGCCTGACACC
TGAAGAAATCGAAAGGATGGTTAATGATGCTGAGAAGTTTGCTGAGGAAG
ACAAAAAGCTCAAGGAGCGCATTGATACTAGAAATGAGTTGGAAAGCTAT
GCCTATTCTCTAAAGAATCAGATTGGAGATAAAGAAAAGCTGGGAGGTAA
ACTTTCTCTGAAGATAAGGAGACCATGGAAAAAGCTGTAGAAGAAAAGA
TTGAATGGCTGGAAAGCCACCAAGATGCTGACATTGAAGACTTCAAAGCT
AAGAAGAAGGAACTGGAAGAAATTGTTCAACCAATTATCAGCAAACTCTA
TGGAAGTGCAGGCCCTCCCCCAACTGGTGAAGAGGATACAGCAGAACTCC
ACCACCACCACCACCAC

SEQ ID NO:128

>gi|21361242|ref|NM_005347.2| Homo sapiens heat shock 70kD protein 5
(glucose-regulated protein, 78kD) (HSPA5), mRNA

ACAGCACAGACAGATTGACCTATTGGGGTGTTCGCGAGTGTGAGAGGGGAAGCGCCGCGGCCTGTATTTC
TAGACCTGCCCTTCGCCTGGTTTCGTGGCGCCTTGTGACCCCGGGCCCTGCCGCTGCAAGTCGGAAATT
GCGCTGTGCTCCTGTGCTACGGCCTGTGGCTGGACTGCCTGCTGCTGCCCAACTGGCTGGCAAGATGAAG

CTCTCCCTGGTGGCCGCGATGCTGCTGCTGCTCAGCGCGGCGCGGGCCGAGGAGGAGGACAAGAAGGAGG
ACGTGGGCACGGTGGTGGCATCGACCTGGGGACCACCTACTCTGCGTGGCGGTTC AAGAACGGCCG
CGTGGAGATCATCGCCAACGATCAGGGCAACCGCATCACGCCGTCTATGTGCGCTTCACTCCTGAAGGG
5 GAACGTCTGATTGGCGATGCCGCCAAGAACCAGCTCACCTCCAACCCCGAGAACACGGTCTTTGACGCCA
AGCGGCTCATCGGCCGACGTGGAATGACCCGTCTGTGCAGCAGGACATCAAGTTCTTGCCGTTCAAGGT
GGTTGAAAAGAAAATAAACCATACATTCAAGTTGATATTTGGAGGTGGGCAAAACAAAGACATTTGCTCCT
GAAGAAATTTCTGCCATGGTTCTCACTAAAATGAAAGAAACCGCTGAGGCTTATTTGGGAAAAGAGTTA
CCCATGCAGTTGTTACTGTACCAGCCTATTTTAATGATGCCCAACGCCAAGCAACCAAAGACGCTGGAAC
10 TATTGCTGGCCTAAATGTTATGAGGATCATCAACGAGCCTACGGCAGCTGCTATTGCTTATGGCTCGGAT
AAGAGGGAGGGGAGAAGAACAATCCTGGTGTGTTGACCTGGGTGGCGGAACCTTCGATGTGTCTCTCTCA
CCATTGACAATGGTGTCTTCGAAGTTGTGGCCACTAATGGAGATACTCATCTGGGTGGAGAAGACTTTGA
CCAGCGTGTCTGGAACACTTCATCAAACGTACAAAAAGAAGACGGGCAAAGATGTCAGGAAAGACAAT
AGAGCTGTGCAGAACTCCGGCGCGAGGTAGAAAAGGCCAAACGGGCCCTGTCTCTCAGCATCAAGCAA
15 GAATTGAAATTGAGTCTTCTATGAAGGAGAAGACTTTTCTGAGACCCTGACTCGGGCCAAATTTGAAGA
GCTCAACATGGATCTGTTCCGGTCTACTATGAAGCCCGTCCAGAAAGTGTGGAAGATTCTGATTTGAAG
AAGTCTGATATTGATGAAAATGTTCTTGTGTTGGTGGCTCGACTCGAATTCCAAAGATTCAAGCACTGGTTA
AAGAGTTCTTCAATGGCAAGGAACCATCCCGTGGCATAAACCCAGATGAAGCTGTAGCGTATGGTGCTGC
TGTCAGGCTGGTGTCTCTGTTGATCAAGATACAGGTGACCTGGTACTGCTTGATGTATGCCCTT
20 ACACCTGGTATTGAAACTGTGGGAGGTGTCTATGACCAAACCTGATTCCAAGGAACACAGTGGTGCCTACCA
AGAAGTCTCAGATCTTTTCTACAGCTTCTGATAATCAACCAACTGTTACAATCAAGGTCTATGAAGGTGA
AAGACCCCTGACAAAAGACAATCATCTTCTGGGTACATTTGATCTGACTGGAATTCCTCCTGCTCCTCGT
GGGGTCCACAGATTGAAGTCACCTTTGAGATAGATGTGAATGGTATTCTTCGAGTGACAGCTGAAGACA
AGGGTACAGGGAACAAAAATAAGATCACAATCACCATGACCAGAATCGCCTGACACCTGAAGAAATCGA
25 AAGGATGGTTAATGATGCTGAGAAGTTTGTGAGGAAGACAAAAGCTCAAGGAGCGCATTGATACTAGA
AATGAGTTGGAAGCTATGCCTATTCTCTAAAGAATCAGATTGGAGATAAAAGAAAAGCTGGGAGGTAAAC
TTTCTCTGAAGATAAGGAGACCATGGAAAAGCTGTAGAAGAAAAGATTGAATGGCTGGAAAGCCACCA
AGATCTGCATGATGAAGACTTCAAAGCTAAGAAGGAAGTGAAGAAATGTTCAACCAATTATCAGC
AAACTCTATGGAAGTGCAGGCCCTCCCCAACCTGGTGAAGAGGATACAGCAGAAAAAGATGAGTTGTAGA
30 CACTGATCTGCTAGTGTCTAATATTGTAAATACTGGACTCAGGAACCTTTTGTAGGAAAAAATTGAAAG
AACTTAAGTCTCGAATGTAATTGGAATCTTACCTCAGAGTGGAGTTGAAACTGCTATAGCCTAAGCGGC
TGTTTACTGCTTTTCATTAGCAGTTGCTCACATGTCTTTGGGTGGGGGGGAGAAGAAGAAATTGGCCATCT
TAAAAAGCAGGTAAAAAACCTGGGTAGGGTGTGTGTTACCTTCAAATGTTCTATTTAACTAAGTGGGT
CATGTGCATCTGGTGTAGGAAGTTTTTCTACCATAAGTGACACCAATAAATGTTTGTATTATTAAGTGG
35 TCTAATGTTTGTGAGAAGCTTCTAATTAGATCAATTACTTATTTTAGGAAATTTAAGACTAGATACTCGT
GTGTGGGGTGAGGGGAGGGAGTATTTGGTATGTTGGGATAAGGAAACACTTCTATTTAATGCTTCCAGGG
ATTTTCTTTTCTTTTAAACCTCTGGGCCCAAGTGATCCTTCCACCTCAGTCTCCAGCTAATTGA
GACCACAGGCTTGTACCACCATGCTCGGCTTTTGCATTAATCTAAGAAAAGGGGAGAGAAGTTAATCCA
CATCTTTACTCAGGCAAGGGGCATTTACAGTGCCCAAGAGTGGGGTTTTCTTGAACATACTTGGTTTCC
TATTTCCCTTATCTTTCTAAACCTGCCTTTCTGGTGGCTTTTTTTAAATATTACTAATGATGCTTTT
40 ATAGCTGCTTGGATTCTCTGAGAAATGATGGGGAGTGAGTGATCACTGGTATTAACCTTATACACTTGG
TTTCATTTGTAACCTTAGGATGTAAAGGTATATTGTGAACCTAGCTGTGTCAGAACTCCATCCCTGAA
ATTTCTCATTAGTGCTACTGGGGTGGGATCTTGGATGGTGACATTGAAACTACACTAAATCCCTCAC'TA
TGAATGGGTTGTTAAAGGCAATGGTTTGTGTCAAACTGGTTTAGGATTACTTAGATTGTGTTCCTGAAG
AAAAGAGTCCAGGTAAATGGTATGATCAATAAAGGACAGGCTGGTGCTAACATAAAATCCAATATTGTAA
45 TCCTAGCACTTTGGGAGGCCAAGGCGGGTGGATCACAAGGTCAAGAGATAGAGACCATCTTTGCCAAT
GGTGAACCTCCATCTCTACTGAAAATACAAAATTAGCTGGGCGTGGTAGTGCAAGCTGAAGGCTGAGGC
AGGAGAATCACTCGAACC CGGGAGGCAGAGGTTGCAGTGAGCCGAGATCACACCACTGTACTCCAGCCCG
GCACTCCAGCCTGGCGACAAGAGTGAGACTCCACCTCAAAAAAAAAAAAAAGAAATCCAATACTGCCCAAG
GATAGGTATTTTATAGATGGGCAACTGGCTGAAAGGTTAATTCTCTAGGGCTAGTAGAACTGGATCCCAA
50 CACCAACTCTTAATTAGACCTAGGCCTCAGCTGCACTGCCCCGAAAAGCATTGGGGCAGACCCTGAGCAG
AATACTGGTCTCAGGCCAAGCCCAATACAGCCATTAAAGATGACCTACAGTGCTGTGTACCTGGGGCAA
TAGGGTTAAATGGTAGTTAGCAACTAGGGCTAGTCTCCCTTACCTCAAAGGCTCTCACTACCGTGGACC
ACCTAGTCTGTAACCTCTTCTGAGGAGCTGTTACTGAATATTAAAAAGATAGACTTCAAAAAAAAAAAAA
AAAAA

SEQ ID NO:129

>gi|16507237|ref|NP_005338.1| heat shock 70kD protein 5 (glucose-regulated protein, 78kD); BiP; Heat-shock 70kD protein-5 (glucose-regulated protein, 78kD) [Homo sapiens]

MKLSLVAAMLLLLSAARAEEDKKEDVGTGVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTP
60 EGERLIGDAKNQLTSNPENTVFDKRLIGRTWNDSVQDQIKFLPFKVVEKTKPYIQVDIGGGQTKTF

APEEISAMVLTKMKETAETGKKVTHAVVTVPAYFNDAQRQATK DAGTIAGLNVMRI INEPTAAAIAYG
LDKREGEKNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGGEDFDQRVMEHF I KLYKKKTGKDVRK
DNRAVQKLRREVEKAKRALSSQHQA RIEIESFYEGEDFSETL TRAKFEELNMDLFRSTMKPVQKVLEDS
LKKSDIDEIVLVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGA AVQAGVLSGDQDTGDLVLLDVC
5 PLTLGIETVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPA
PRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKLKERID
TRNELESYAYS LKNQIGDKELGGKLSSSEDKETMEKAVEEKI EWLESHQDADIEDFKAKKKELEEIVQPI
ISKLYGSAGPPPTGEEDTAEKDEL

10 SEQ ID NO:130

IFNgR1

>G3_2_37_PCR_G3F1 Direction: sense

CTGGCTTTAACTCTGACCCAAAGTAGTATTTGATGGATCACCAACATGTATCAGTAAATATTACAATAATGNATG
TAGTAATATTGATGCAGGCATCAATCCATTCTGAATTCCTAACACCATAGTTCTTTACCTCTACGGTAAAAACAG
15 GGACCTGTGGCATGATCTGGTACTCCCAATATACGATAGGGTTCATGTTATAGGATTCAATTGTAACATTAGTTG
GTGTAGGCACTGAGGACGGCCCCAGATCCGCGGTGCCATCTCAGCCCTGCTCACACCCTGCATGACAAGGGGTA
GGAGAAAGAGGAGGCCATGCTGCTACCGACGGTCGCTGGCTCCAACCCCGAGCGCCTGCGGGACCAGCCAGCG
CTGCCC

20 SEQ ID NO:131

>GL3_2.37_A_M13F Direction: sense

CTGGCTTTAACTCTGACCCAAAGAGAATTTGATGGATCACCAACATGTATCAGAAATATTACAATAATGATGAGAA
ATATTGATGCAGGCATCAATCCATTCTGAATTCCTAACACCATAGTTCTTTACCTCTACGGTAAAAACAGGGACC
TGTGGCATGATCTGGTACTCCCAATATACGATAGGGTTCATGTTATAGGATTCAATTGTAACATTAGTTGGTGT
25 GGCATGAGGACGGCCCCAGATCCGCGGTGCCATCTCAGCCCTGCTCACACCCTGCATGACAAGGGGTAGGAGA
AAGAGGAGAGCCATGCTGCTACCGACGGTCGCTGGCTCCAACCCCGAGCGCCTGCGGGACCAGCCAGCGCTGCC
C

SEQ ID NO:132

>GL3_2.37_A_M13R Direction: anti-sense

GGGCAGCGCTGGGCTGGTCCCCGAGGCGCTCGGGGTTGGAGCCAGCGACCGTCGGTAGCAGCATGGCTCTCCTCT
TTCTCCTACCCCTTGTATGCAGGGTGTGAGCAGGGCTGAGATGGGCACCGCGGATCTGGGGCCGCTCCTCAGTGC
CTACACCAACTAATGTTACAATTGAATCCTATAACATGAACCCCTATCGTATATTGGGAGTACCAGATCATGCCAC
AGGTCCCTGTTTTTACCGTAGAGGTAAAGAACTATGGTGTTAAGAATTGAGATGGATTGATGCCTGCATCAATA
35 TTTCTCATCATTATTGTAATATTTCTGATCATGTTGGTGATCCATCAAATTCTCTTTGGGTCAGAGTTAAAGCCA
G

SEQ ID NO:133

>GL3_2.37_B_M13F Direction: anti-sense

GGGCAGCGCTGGGCTGGTCCCCGAGGCGCTCGGGGTTGGAGCCAGCGACCGTCGGTAGCAGCATGGCTCTCCTCT
TTCTCCTACCCCTTGTATGCAGGGTGTGAGCAGGGCTGAGATGGGCACCGCGGATCTGGGGCCGCTCCTCAGTGC
CTACACCAACTAATGTTACAATTGAATCCTATAACATGAACCCCTATCGTATATTGGGAGTACCAGATCATGCCAC
AGGTCCCTGTTTTTACCGTAGAGGTAAAGAACTATGGTGTTAAGAATTGAGATGGATTGATGCCTGCATCAATA
45 TTTCTCATCATTATTGTAATATTTCTGATCATGTTGGTGATCCATCAAATTCTCTTTGGGTCAGAGTTAAAGCCA
G

SEQ ID NO:134

>GL3_2.37_B_M13R Direction: sense

CTGGCTTTAACTCTGACCCAAAGAGAATTTGATGGATCACCAACATGTATCAGAAATATTACAATAATGATGAGAA
ATATTGATGCAGGCATCAATCCATTCTGAATTCCTAACACCATAGTTCTTTACCTCTACGGTAAAAACAGGGACC
TGTGGCATGATCTGGTACTCCCAATATACGATAGGGTTCATGTTATAGGATTCAATTGTAACATTAGTTGGTGT
50 GGCATGAGGACGGCCCCAGATCCGCGGTGCCATCTCAGCCCTGCTCACACCCTGCATGACAAGGGGTAGGAGA
AAGAGGAGAGCCATGCTGCTACCGACGGTCGCTGGCTCCAACCCCGAGCGCCTGCGGGACCAGCCAGCGCTGCC
C

SEQ ID NO:135

>GL3_2.37_C_M13F Direction: anti-sense

GGGCAGCGCTGGGCTGGTCCCCGAGGCGCTCGGGGTTGGAGCCAGCGACCGTCGGTAGCAGCATGGCTCTCCTCT
TTCTCCTACCCCTTGTATGCAGGGTGTGAGCAGGGCTGAGATGGGCACCGCGGATCTGGGGCCGCTCCTCAGTGC
CTACACCAACTAATGTTACAATTGAATCCTATAACATGAACCCCTATCGTATATTGGGAGTACCAGATCATGCCAC
60

AGGTCCTGTTTTTACCGTAGAGGTAAAGAACTATGGTGTTAAGAATTCAGAATGGATTGATGCCTGCATCAATA
TTTCTCATCATTATTGTAATATTTCTGATCATGTTGGTGATCCATCAAATTCCTTTGGGTGAGAGTTAAAGCCA
G

5 SEQ ID NO:136

>GL3_2.37_C_M13R Direction: sense

CTGGCTTTAACTCTGACCCAAAGAGAATTTGATGGATCACCAACATGATCAGAAATATTACAATAATGATGAGAA
ATATTGATGCAGGCATCAATCCATTCTGAATTCCTTAACACCATAGTTCTTTACCTCTACGGTAAAAACAGGGACC
TGTGGCATGATCTGGTACTCCCAATATACGATAGGGTTCATGTTATAGGATTCAATTGTAACATTAGTTGGTGTA
10 GGCATGAGGACGGCCCCAGATCCGCGGTGCCATCTCAGCCCTGCTCACACCTGTCATGACAAGGGGTAGGAGA
AAGAGGAGAGCCATGCTGCTACCGACGGTGCCTGGCTCCAACCCGAGCGCCTGCGGGACCAGCCCAGCGCTGCC
C

SEQ ID NO:137

>GL3_2.37_D_M13F Direction: sense

CTGGCTTTAACTCTGACCCAAAGAGAATTTGATGGATCACCAACATGATCAGAAATATTACAATAATGATGAGAA
ATATTGATGCAGGCATCAATCCATTCTGAATTCCTTAACACCATAGTTCTTTACCTCTACGGTAAAAACAGGGACC
TGTGGCATGATCTGGTACTCCCAATATACGATAGGGTTCATGTTATAGGATTCAATTGTAACATTAGTTGGTGTA
GGCATGAGGACGGCCCCAGATCCGCGGTGCCATCTCAGCCCTGCTCACACCTGTCATGACAAGGGGTAGGAGA
20 AAGAGGAGAGCCATGCTGCTACCGACGGTGCCTGGCTCCAACCCGAGCGCCTGCGGGACCAGCCCAGCGCTGCC
C

SEQ ID NO:138

>GL3_2.37_D_M13R Direction: anti-sense

GGGCAGCGCTGGGCTGGTCCCGCAGGCGCTCGGGGTGGAGCCAGCGACCGTCGGTAGCAGCATGGCTCTCCTCT
TTCTCCTACCCCTTGTCTATGCAGGGTGTGAGCAGGGCTGAGATGGGCACCGCGGATCTGGGGCCGTCCTCAGTGC
CTACACCAACTAATGTTACAATTGAATCCTATAACATGAACCTATCGTATATTGGGAGTACCAGATCATGCCAC
AGGTCCTGTTTTTACCGTAGAGGTAAAGAACTATGGTGTTAAGAATTCAGAATGGATTGATGCCTGCATCAATA
TTTCTCATCATTATTGTAATATTTCTGATCATGTTGGTGATCCATCAAATTCCTTTGGGTGAGAGTTAAAGCCA
30 G

SEQ ID NO:139

>gi|4557879|ref|NM_000416.1| Homo sapiens interferon gamma receptor 1
(IFNGR1), mRNA

CCGCAGGCGCTCGGGGTGGAGCCAGCGACCGTCGGTAGCAGCATGGCTCTCCTCTTTCTCCTACCCCTT
GTCATGCAGGGTGTGAGCAGGGCTGAGATGGGCACCGCGGATCTGGGGCCGTCCTCAGTGCCTACACCAA
CTAATGTTACAATTGAATCCTATAACATGAACCTATCGTATATTGGGAGTACCAGATCATGCCACAGGT
CCCTGTTTTTACCGTAGAGGTAAAGAACTATGGTGTTAAGAATTCAGAATGGATTGATGCCTGCATCAAT
ATTTCTCATCATTATTGTAATATTTCTGATCATGTTGGTGATCCATCAAATTCCTTTGGGTGAGAGTTA
40 AAGCCAGGGTGGACAAAAAGAATCTGCCTATGCAAAGTCAGAAGAATTTGCTGTATGCCGAGATGGAAA
AATTGGACCACCTAAACTGGATATCAGAAAGGAGGAGAAGCAAAATCATGATTGACATATTTACCCCTTCA
GTTTTTGTAAATGGAGACGAGCAGGAAGTCGATTATGATCCCCGAACTACCTGTTACATTAGGGTGTA
ATGTGTATGTGAGAATGAACGGAAGTGAGATCCAGTATAAAATACTCACGCAGAAGGAAGATGATTGTGA
CGAGATTCACTGCCAGTTAGCGATTCCAGTATCCTCACTGAATTCTCAGTACTGTGTTTCAGCAGAAGGA
45 GTCTTACATGTGTGGGTGTTACAACCTGAAAAGTCAAAGAAGTTTGTATTACCATTTTCAATAGCAGTA
TAAAAGGTTCTCTTTGGATTCCAGTTGTTGCTGCTTTACTACTCTTTCTAGTGCTTAGCCTGGTATTCAT
CTGTTTTTATATTAAGAAAATTAATCCATTGAAGGAAAAAAGCATAATATTACCCAAGTCCTTGATCTCT
GTGGTAAGAAGTGCTACTTTAGAGACAAAACCTGAATCAAAATATGTATCACTCATCACGTCATACCAGC
CATTTTCCTTAGAAAAGGAGGTGGTCTGTGAAGAGCCGTTGTCTCCAGCAACAGTTCCAGGCATGCATAC
50 CGAAGACAATCCAGGAAAAGTGGAACATACAGAAGAACTTTCTAGTATAACAGAAGTGGTGACTACTGAA
GAAAATATTCCTGACGTGGTCCCGGGCAGCCATCTGACTCCAATAGAGAGAGAGAGTTCTTACACCTTTAA
GTAGTAACCACTCTGAACCTGGCAGCATCGCTTTAACTCGTATCACTCCAGAAATTGTTCTGAGAGTGA
TCACTCCAGAAATGGTTTTGATACTGATTCCAGCTGTCTGGAATCACATAGCTCCTTATCTGACTCAGAA
TTTCCCCCAAATAATAAAGGTGAAATAAAAAACAGAAGGACAAGAGCTCATAACCGTAATAAAAGCCCCCA
55 CCTCCTTTGGTTATGATAAACCACATGTGCTAGTGGATCTACTTGTGGATGATAGCGGTAAAGAGTCCTT
GATTGGTTATAGACCAACAGAAGATTCCAAAGAATTTTCATGAGATCAGCTAAGTTGCACCAACTTTGAA
GTCTGATTTTCTGGACAGTTTTCTGCTTTAATTTTCATGAAAAGATTATGATCTCAGAAATTGTATCTTA
GTTGGTATCAACCAAATGGAGTGACTTAGTGATCATGAAAGCGTAAAGAGGATGTGTGGCATTTTCACTT
TTGGCTTTGTAAGTACAGACTTTTTTTTTTTTTTAAACAAAAAAGCATTGTAACCTTATGAACCTTTACA
60 TCCAGATAGGTTACCAGTAACGGAACATATCCAGTACTCCTGGTTCCTAGGTGAGCAGGTGATGCCCCAG

GGACCTTTGTAGCCACTTCACTTTTCTCTCTGCCTTGGTATAGCATATGTTTGTAAAGTTTA
TGCATACAGTAATTTTAAGTAATTTTCAGAAGAAATTCTCGAAGCTTTTCAAATTTGGACTTAAAATCTAA
TTCAAATAATAGAATTAATGGAATATGTAAATAGAAACGTGTATATTTTTTATGAAACATTACAGTTAG
AGATTTTTAAATAAAGAATTTTAAAACTC

SEQ ID NO:140

>gi|4557880|ref|NP_000407.1| interferon gamma receptor 1; Immune
interferon, receptor for [Homo sapiens]

MALLFLLPLVMQGVSRAMGTADLGPSSVPTPTNVITIESYNNMPIVYWEYQIMPQVPVFTVEVKNYGVKN
SEWIDACINISHYCNISDHVGDPSNSLWVRVKARVGQKESAYAKSEEFVAVCRDGKIGPPKLDIRKEEKQ
IMIDIFHPSVFNVDQEVDDPETTCYIRVYNVYVRMNGSEIQYKILTQKEDDCDEIQCQLAIPVSSLN
SQYCVSAEGLVHVWGVTTTEKSKVCITIFNSSIKGSLWIPVVAALLLFLVLSLVFICFYIKKINPLKEKS
IILPKSLISVVRSATLETKEPKYVSLITSYQPFSLKEVCEPLSPATVPGMHTEDNPGKVEHTEELS
SITEVVTTEENIPDVVPGSHLTPIERESSPLSSNQSEPGSIALNSYHSRNCSES DHSRNGFDTDSSCLE
SHSSLS DSEFPNNKGEIKTEGQELITVIKAPTSTFGYDKPHVLVDLLVDDSGKESLIGYRPTEDSKEFS

SEQ ID NO:141

Importin a4

>GL2_91_2_M13F Direction: sense

CTATTGTGCTTGCTTCATCACCGGCCATTATCAGAATGTTTTTTAGACCATCTAGAACCACCTGAACCACCTTGAG
AATCTTTCACTGACAGTAAATTACAGAACGGTGGTATTACATTCTGCTGTACAAGGTAAGTCAACCTGATCTTTTC
TGCCACTTATTGTTAAGTTGCTGATTGCCCAAGCAGCTTCTTTTGTGTTCCAAAGTCCCCCTTAGCAAGCTGAT
GAATTATCATAGGAATTAATCCAGCATCTATTACAGCTTGAACCTTGTGCTGGTTGCCTGCTGTTATGTTGGAAA
GGAACCACACTGCTTCCTTATTTATCTTCTCTTTTGGGTGTGATAAGAGATTTGGGAAGTGTGACAGGACATCAC
AATTGAGAACAACTGGGTCTGCTCGTCCGTGCCAGTCACTATGTTGCCAACTGCTCTGAGGGCTGCTGTTTGAA
CTTTGACTTCCTGATGGCTCAGAAGGGGCACAAGAAAGGGCACAACCTCCTGAATTCCAATAACCATCTGGTATCT
GTTTATTACCTTCCATCTGTCAAAGTTATGACAGAGCCCAAAACAAGTGTCTTACAAGAAATGTTTATATCTGGA
ATGGGTAATATGAAGGACAACATAAAAGGC

SEQ ID NO:142

>GL2_91_1_M13R Direction: sense

CTATTGTGCTTGCTTCATCACCGGCCATTATCAGAATGTTTTTTAGACCATCTAGAACCACCTGAACCACCTTGAG
AATCTTTCACTGACAGTAAATTACAGAACGGTGGTATTACATTCTGCTGTACAAGGTAAGTCAACCTGATCTTTTC
TGCCACTTATTGTTAAGTTGCTGATTGCCCAAGCAGCTTCTTTTGTGTTCCAAAGTCCCCCTTAGCAAGCTGAT
GAATTATCATAGGAATTAATCCAGCATCTATTACAGCTTGAACCTTGTGCTGGTTGCCTGCTGTTATGTTGGAAA
GGAACCACACATGCTTCCTTATTTATCTTCTCTTTAGGGTGTGATAAGAGATTTGGGAAGTGTGACAGGACATCA
CAATTTGAGGAACAACCTGGGTCTGACTCGTCCGTGCCAGTCACTATGTTGCCAACTGCATCTGAGGGCCTGCTG
TTTGAAACTTTGACTTCCTGATGGCTCAGAAGGGGCACAAGAAAGGGCNACAACTCCTGGAATCNCAATAAAC
AATCTGTATCTGTTCAATAACCTCCATCCTGGTGCAAAGGTANCTGAACACGGGAGGCCCAAAACAGGTGGT
CCTACAAGGAAATGGTGAACAAACATGGAAATGGGAATAGGANGGACCACCACATAAAGG

SEQ ID NO:143

>GL2_91_3_M13R Direction: sense

CTATTGTGCTTGCTTCATCACCGGCCATTATCAGAATGTTTTTTAGACCATCTAGAACCACCTGAACCACCTTGAG
AATCTTTCACTGACAGTAAATTACAGAACGGTGGTATTACATTCTGCTGTACGAGGTAAGTCAACCTGATCTTTTC
TGCCACTTATTGTTAAGTTGCTGATTGCCCAAGCAGCTTCTTTTGTGTTCCAAAGTCCCCCTTAGCAAGCTGAT
GAATTATCATAGGAATTAATCCAGCATCTATTACAGCTTGAACCTTGTGCTGGTTGCCTGCTGTTATGTTGGAAA
GGAACCACACTGCTTCCTTATTTATCTTCTCTTTGGGTGTGATAACGAGATTTGGGAAGTGTGACAGGACATCA
CAATTGAGAACAACTGGGTCTGCTCGTCCGTGCCAGTCACTATGTTGCCAACTGCTCTGAGGGCTGCTGTTTGT
AAGTTGACTTCCTGATGGCTCAGCAAGGGGCACAAGAAAGGGCACAACCTCCCTGAAAATCCAATAAACCAA
TCTGTAATTTCTGGTTCACATTTACCTTCCAATTTCTGGTTCAAAGGTATGACAGAAGCCCCAAAACAGGTGTTCT
ACCAAGAAATTTGTTACATAATCCTGAATGGGAATANTGACNNGGACACACATTAAAGCCGCAAAAACCANACANA
CAAACACAGAC

SEQ ID NO:144

>GL2_202_1_M13R Direction: anti-sense

CTTTATGTGTCCTCATATACCATAACAGATATAAACATTCTTGTAGACACTGTTTGGGCTCTGTGCATCTTGACAG
ATGGAGGTAATGAACAGATACAGATGGTTATTGATTCAGAGATTGTGCCCTTTCTTGATGCCCTTCTTTGGGGA
AAANCCCCAACTCCACGGGGGAGAAGCGTCCNAACAAAGNGTTCCCCACAAAACAAGCACCACAACGCGGCCG
AAGGACACCGCAGTTCCAACNGGAAAAAGAGGAGCCCAACACACGGGAGCATAATATNTGAGGGGGGGGCC

AAGAAAACACCCGCAACTTAAAANGCGCNTGGGGAAGAGCCATTAGGTGGGNNCAACCGCCCCGGGAAAAGCCGC
GGAAAAGGGCCCAANGGAAAAACACCCACCCCANCCCCAANAACACGGGGGGGAGAATGCATCACGGGGCGAG
ATCAATACCAGATTTACCCAAAAACATTCACAACG

5 SEQ ID NO:145

>GL2_202_1_M13F Direction: sense

CTATTGTGCTTGCTTCATCACCGGCCATTATCAGAATGTTTTTTAGACCATCTAGAACCACCTGAACCACTTGAG
AATCTTTTCACTGACAGTAAATTACAGAACGGTGGTATTACATTCTGCTGGTTTAACAAGGTACTCAACCTGATCT
TTTCTGCCACTTATTGTTAAGTTGCTGATTGCCCAAGCAGCTTCTTTTGTGTTCCAAAGTCCCCCTTAGCAAGG
10 CTGATGAATTATCATAGGAATTAATCCAG

SEQ ID NO:146

>GL2_66_F10_G3F1 Direction: N/A

CTATTGTGCTTGCTTCATCACCGGCCATTATCAGAATGTTTTTTAGACCATCTAGAANCACCTGAACCACTTGAG
15 AATCTTTTCACTGACAGTAAATTACAGAACGGTGGTATTACATTCTGCTGTACAAGGGGACTCAACTGATCTTTTC
TGGCACTTATTGGTAAGTTGCTGATTGGCCAAGGCAGCTTCTTTTGTGTTCCAAAGTTCCCTTAGCAAGCTGA
TGAATTATCATAGGAATTAATCCAGCATCTATTACAGCTTGGNACCTGTTGCAGGGTTGACCTGCTGGATAATGT
TTGGGAAGGAAACACACTGGGTTCCTTAATTTATCTTATCTCTATAGGGAGTGATAAGAGATATATGGGAAAGT
20 TGTTAACAGGACATTACAAATTTGATGAACAACCTGGGTCTGATACGTCNCAGNTGGGCCACNCGCAACGCAAA
TAAGTTTTGGGCCCAACCAGGGGGCCACGAAAAAGAGGACGCAGCCACGTTGGNTTGGAGAACCATCTTGA
AACANTACCCTCNGGAATAGAGCGTCCCAANAGGGGGCGGACCCAAGGAANAAAGGGGGACCAAAGTACCCTGG
AAATCCACACATAAAACCATACTGTAATACTCGTGTGCCACGCTGTTAATACCCANTTCACCAACAATCCACTG
GGGTCNACACAGGATNANTTGAAAGCACACGGAAGAGAGCCGCCAACAAACCAGNGGGGGTGCGTNTGTTGCCAAG
25 CAACCACANGGGTGCGCTGAAGTAGCAATAACCAGGANAANTGGGGGGACGAGCAACTTNGTAAAGGGTGAACG

SEQ ID NO:147

>GL2_91_1_M13F Direction: anti-sense

CTTTATGTGTCCTCATATACCATACAGATATAAACATTCTTGTAGACACTGTTTGGGCTCTGTCTACTTGTACAG
ATGGAGGTAATGAACAGATACAGATGGTTATTGATTAGGAGTTGTGCCCTTTCTTGTGCCCTTCTGAGCCATC
30 AGGAAGTCAAAGTTCAAACAGCAGCCCTCAGAGCAGTTGGCAACATAGTGACTGGCACCGACGAGCAGACCCAGG
TTGTTCTCAATTGTGATGTCTGTCTCACACTTCCCAAATCTCTTATCACACCCAAAAGAGAAGATAAATAAGGAAG
CAGTAGTGGTTCCCTTCCAACATAACAGCAGGCAANCAGCAACAAGTTCAAGCTGTAATAGATGCTGGATTAATT
CCTATGATAATTTATCAGCNTTGCTAAGNGGGGACATATTGGAACACAAAAAGAAGCTGCATATGGGGCAAC
35 AATCAGGCAANAC

SEQ ID NO:148

>GL2_91_2_M13R Direction: anti-sense

CTTTATGTGTCCTCATATACCATACAGATATAAACATTCTTGTAGACACTGTTTGGGCTCTGTCTACTTGTACAG
ATGGAGGTAATGAACAGATACAGATGGTTATTGATTAGGAGTTGTGCCCTTTCTTGTGCCCTTCTGAGCCATC
40 AGGAAGTCAAAGTTCAAACAGCAGCCCTCAGAGCAGTTGGCAACATAGTGACTGGCACCGACGAGCAGACCCAGG
TTGTTCTCAATTGTGATGTCTGTCTCACACTTCCCAAATCTCTTATCACACCCAAAAGAGAAGATAAATAAGGAAG
CAGTGTGGTTCCCTTCCAACATAACAGCAGGCAACCAGCAACAAGTTCAAGCTGTAATAGATGCTGGATTAATTC
CTATGATAATTCATCAGCTTGCTAAGGGGGACTTTGGAACACAAAAAGAAGCTGCTTGGGCAATCAGCAACTTAA
45 CAATAAGTGGCAAGAAAAGATCAGGTTGAGTACCTTGTAACAGCAGAATTGTAAATACCACCGTTCTGTAAATTT

SEQ ID NO:149

>GL2_91_3_M13F Direction: anti-sense

CTTTATGTGTCCTCATATACCATACAGATATAAACATTCTTGTAGACACTGTTTGGGCTCTGTCTACTTGTACAG
ATGGAGGTAATGAACAGATACAGATGGTTATTGATTAGGAGTTGTGCCCTTTCTTGTGCCCTTCTGAGCCATC
50 AGGAAGTCAAAGTTCAAACAGCAGCCCTCAGAGCAGTTGGCAACATAGTGACTGGCACCGACGAGCAGACCCAGG
TTGTTCTCAATTGTGATGTCTGTCTCACACTTCCCAAATCTCTTATCACACCCAAAAGAGAAGATAAATAAGGAAG
CAGTGTGGTTCCCTTCCAACATAACAGCAGGCAACCAGCAACAAGTTCAAGCTGTAATAGATGCTGGATTAATTC
CTATGATAATTCATCAGCTTGCTAAGAGGGGACTTTGGAACACAAAAAGAAGCTGCTTGGGCAATCAGCAACTT
AACAAATAAAGTGGCAGAAAAGATCAGGTTGAGTACCTCGTACAGCAGAATTGTAAATACCACCGATACTGTTAACT
55 TNTACCTGTCTAGTGAACAGATCCTCAAGTGGATCANAGAGGGACCTAGAATGGGTCTAAAAAACAATNTC
TGGATAAATTTNGGGCCGGGAGGACATGAAAAGCAAAGGCAACAAATCAAGGACACGGAACGAGAAACCAAGCANA
CACATAGGAAGACCACCCGNTGAGCCAACCCGGGACTATGAGGAGGGGCGTCCCTCCATGGANACAGGAGCAGC
GACCCAAGNCCAAAAAAGGGGCGAATACAGTAGAACAAGAGNAGNAACAACACCATAACAAGAGAGAANAACACC
60 GACAGCCA

SEQ ID NO:150

>GL2_66_F07_G3F1 Direction: N/A

CTATTGTGCTTGCTTCATCACCGGCCATTATCAGAATGTTTTTTAGACCATCTAGAACCACCTGAACCACTTGAG
AATCTTTCACTGACAGTAAATTACAGAACGGTGGTATTACATTCTGCTGTACAAGGTACTCAACCTGATCTTTTC
5 TGCCACTTATTGTTAAGTTGCTGATTGCCCAAGCAGCTTCTTTTTGTGTTCCAAAGTCCCCCTTAGCAAGCTGAT
GAATTATCATAGGAATTAATCCAGCATCTATTACAGCTTGAACCTGTTGCTGGTTGCCTGCTGTTATGTTGGAAA
GGAACCACACTGCTTCCTTATTTATCTTCTCTTTTGGGTGTGATAAGAGATTTGGGAAGTGTGACAGGACATCAC
AATTGAGAACAACCTGGGTCTGCTCGTCCGTGCCAGTCACATGTTGCCAACTGCTCTGAGGGCTGCTGTTTGAA
10 CTTTGACTTCCTGATGGCTCAGAAGGGGCACAAGAAAGGGCACAACCTCCTGAATCAATAACCATCTGTATCTGTT
CATTACCTCCATCTGTCAAGTATGACAGAGCC

SEQ ID NO:151

>GL2_66_F08_G3F1 Direction: N/A

CTATTGTGCTTGCTTCATCACCGGCCATTATCAGAATGTTTTTTAGACCATCTAGAACCACCTGAACCACTTGAG
15 AATCTTTCACTGACAGTAAATTACAGAACGGTGGTATTACATTCTGCTGTACAAGGTACTCAACCTGATCTTTTC
TACCACTTATTGTTAAGTTGCTGATTGCCCAAGCAGCTTCTTTTTGTGTTCCAAAGTCCCCCTTAGCAAGCTGAT
GAATTATCATAGGAATTAATCCAGCATCTATTACAGCTTGAACCTGTTGCTGGTTGCCTGCTGTTATGTTGGAAA
GGAACCACACTGCTTCCTTATTTATCTTCTCTTTTGGGTGTGATAAGAGATTTGGGAAGTGTGACAGGACATCAC
AATTGAGAACAACCTGGGTCTGCTCGTCCGTGCCAGTCACATGTTGCCAACTGCTCTGAGGGCTGCTGTTTGAA
20 CTTTGACTTCCTGATGGCTCAGAAGGGGCACAAGAAAGGGCACAACCTCCTGAATCAATAACCATCTGTATCTGTT
CATTACCTCCATCTGTCAAGTATGACAGAAGCCCAAACAGCTGTCTACAAGAATGTTTATATCTGTATGGTATAT
TGAAGGACACCATAAAG

SEQ ID NO:152

>GL2_66_F09_G3F1 Direction: N/A

CTATTGTGCTTGCTTCATCACCGGGCCATTATCAGAATGTTTTTTAGACCATCTAGAACCACCTGAACCACTTG
AGAATCTTTCACTGACAGTAAATTACAGAACGGTGGTATTACATTCTGCTGTACAAGGTACTCAACCTGATCTTT
TCTGCCACTTATTGTTAAGTTGCTGATTGCCCAAGCAGCTTCTTTTTGTGTTCCAAAGTCCCCCTTAGCAAGCTG
30 ATGAATTATCATAGGAATTAATCCAGCATCTATTACAGCTTGAACCTGGTTGCTGGTTGCCTGCTGTTATGTTGG
AAAGGAACCACACTGCTTCCTTATTTATCTTCTCTTTTGGGTGTGATAAGAGATTTGGGAAGTGTGACAGGACAT
CACAATTGAGAACAACCTGGGTCTGCTCGTCCGTGCCAGTCACATGTTGCCAACTGCTCTGANGGCTGCTGTTTG
AACTTTGACTTCCTGATGGCTCAAAGGGGCACAAGAAAGGGCACAACCTCCTGAATCAATAACCATCTGTATCTGT
GTCATTACCTCCATCTGTCAAGTATGACAGAGCCCAAACAGTGTCTACACAGCAATGTTCTATATCTCGTATGGN
35 GATATGAGGGACACCTNAAAGCCG

SEQ ID NO:153

>gi|4504898|ref|NM_002267.1| Homo sapiens karyopherin alpha 3 (importin
alpha 4) (KPNA3), mRNA

AATTCGTGGTTCCAGAATCGATAGGGCTCCAAGATTGCGCCGCCGCCGCCGCCGAGCCGCAGGAGTAGCC
40 GCCGCCGAGCCGCGCGCAGCATGGCCGAGAACCCAGCTTGGAGAACCACCGCATCAAGAGCTTCAAGA
ACAAGGGCCGCGATGTGGAACAATGCGAAGACATAGAAATGAAGTGACAGTGGAACCTGCGGAAGAACAA
AAGAGATGAACACTTATTGAAAAAGAGAAATGTTCCCAAGAAGAAAGTCTAGAAGATTCAGATGTTGAT
GCTGATTTTAAAGCACAAAATGTAACCTTAGAAGCTATATTGCAGAATGCCACAAGTGATAACCCGATTGG
TCCAATTGAGTGCTGTCCAGGCAGCAAGAAAACCTGTTATCCAGTGACCAGAATCCACCGATTGATGACTT
45 AATAAAATCTGGGATTTTACCAATTCTAGTCAAATGTCTAGAAAGGGATGATAATCCTTCATTACAGTTT
GAAGCTGCTTGGGCATTAACATAAGCATCAGGAACCTCTGCACAGACTCAAGCTGTTGTGCAGTCTA
ATGCAGTACCTCTTTTTCTGAGACTTCTTCGTTCCACCACATCAGAATGTTTGTGAACAAGCAGTATGGGC
TTTGGGAAACATTATAGGTGATGGTCCTCAATGTAGAGATTATGTATATCACTGGGAGTTGTCAAACCT
CTTCTGTCCTTCATCAGTCCCTCCATCCCATCACCCTTCGGAACGTCACATGGGTCATTGTCAATC
50 TCTGCAGGAATAAGGATCCCCACCGCTATGGAGACAGTTTCAAGGATTTTGGCAGCTTTATGTGTCCT
CATATACCATAACAGATATAAACATTCTGTAGACACTGTTTGGGCTCTGTCACTTGACAGATGGAGGT
AATGAACAGATACAGATGGTTATTGATTGAGGAGTTGTGCCCTTCTGTGCCCCCTCTGAGCCATCAGG
AAGTCAAAGTTTCAAACAGCAGCCCTCAGAGCAGTTGGCAACATAGTGACTGGCACCAGCAGCAGACCCA
GGTTGTTCTCAATTGTGATGTCCTGTACACTTCCCAAATCTCTTATCACACCCAAAAGAGAAGATAAAT
55 AAGGAAGCAGTGTGGTTCTTTTCCAACATAACAGCAGGCAACCAGCAACAAGTTCAAGCTGTAATAGATG
CTGGATTAATTCCTATGATAATTCATCAGCTTGCTAAGGGGGACTTTGGAACACAAAAGAGCTGCTTG
GGCAATCAGCAACTTAACAATAAGTGGCAGAAAAGATCAGGTTGAGTACCTTGTACAGCAGAATGTAATA
CCACCGTTCTGTAATTTACTGTGAGTGAAGATTCTCAAGTGGTTTCAAGTGGTTCTAGATGGTCTAAAAA
ACATTCTGATAATGGCCGGTGATGAAGCAAGCACAATAGCTGAAATAATAGAGGAATGTGGAGGTTTGA
60 GAAAATTGAAGTTTTACAGCAACATGAAAATGAAGACATATATAAATTAGCATTGAAATCATAGATCAG

TATTTCTCTGGTGATGATATTGATGAAGATCCCTGCCTCATTCTCTGAAGCAACACAAGGAGGTACCTACA
ATTTTGATCCAACAGCCAACCTTCAAACAAAAGAATTTAATTTTAAATTCAGTTGAGTGCAGCATCTTT
CCCACATTCAATATGAAGCACCACAGATGGCTACCAAATGATAAGAACAACAGCAACAAAAGGCTCCAA
AACACACATGCCTCTTTGTTTTGATGCTTCTAAAGAAAGCCATGTCTCAGTCACTTTGCAGTTGCCAAAA
5 GTCATATCACATGGACTGTAAATGCATATGCATGATTTCTAAACTGTTTTAGAACTCTCCTTAACAAT
CTCAACTACCTATTTTTCCCTGTTCCCTGGTGCCACAGGCTGACAACTGCAGTCTCCAGTTTAGAATAA
ATATTCCATAGTGGTGACATGTCAGCTGCCCACTGATACTCTTTGGAAAATGGTGCCTGTGGATCAAG
ACACTTTGGTATGATGCATATACAAGTTGGAAGACTAAAGAGGTGCAATGTGATCTGAGCCTCCATCATT
10 GTCCTCCACAAACATATTTTCATATTCCTTATGTGGAAGAATAGATTTTAAAGTACAAGCCAAATGATTT
TCATTGGTGGAAGTACACAAAAAAGTAACTTAAAAACAAGAACTTGTTATTGAATAAACAGATAA
GTTTT

SEQ ID NO:154

>gi|4504899|ref|NP_002258.1| karyopherin alpha 3; importin alpha 4 [Homo sapiens]

MAENPSLENHRIKSFKNKGRDVETMRRHRNEVTVELRKNKRDEHLLKKRNVPQEESSLESDVDADFKAQN
VTLEAILQNASDNPVQLSAVQAARKLLSSDQNPIDDLIKSGILPILVKCLERDDNPSLQFEAAWALT
NIASGTSAQTAQVQSNAPVPLFLRLLRSPHQNVCEQAVWALGNIIGDGPQCRDYVISLGVVKPLLSFISP
SIPITFLRNVTWVIVNLCRNKDPPTMETVQEILPALCVLIYHTDINILVDTVWALSYLETDGGNEQIQMV
20 IDSGVVPFLVPLLSHQEVKVQTAALRAVGNIVTGTDEQTQVVLNCDVLSHFNNLLSHPKEKINKEAVWFL
SNITAGNQOQVQAVIDAGLIPMIHQIAKGDFTGQKEAAWAIISNLTISGRKDQVEYLVQQNVIPFCNLL
SVKDSQVVQVVLVDGLKNILIMAGDEASTIAEIIIECGGLEKIEVLQGHENEDIYKLAFEIIDQYFSGDDI
DEDPLCIPEATQGGTYNFDPTANLQTKENF

SEQ ID NO:155

Lysosomal pepstatin-insensitive protease (CLN2)

>GL3_21A_2_M13F Direction: N/A

CCATTTCAATACTATGTGGTCGTTACAATGCTAGAGGTAGTGACAGGGGGCGCAATGCAAGGGAGGAAGGGCGT
TAACATCTGCCACCTACTTCCAGGTGCCAAGCACTGTTATCAACATTATTCCACTTTATTCCCTGGGTGATTAT
30 GAAAGGCCAGGTATTGAGTATTCACACTTNACAGACGAGGAAAACAAGCGCTCAGGGNAGAATAAGCGTTACTTG
ANCCAGGTACATACAATACGTTACCTAGAGTCCANTTATTTCTAGGAAACCACAGGAGTTTCCAAAACAGGGGTG
TTTGGTGTNAGGAAACAATCNTATGGGGAGATATTTTTCCAAGCGCAATTACCACAACATTATAAAGTANGCCT
CCAGGGGGGAAAATCAAAACCATTCACGAAAAGAAAGGTNGAACATTCTGGAGAGCATTGAAGGATCTTTGCG
CAAAGAAATTCCTTGGGAAAAATTTACCTTGGGCAAGGTATGTTATCAAGAGGGGTTGGGTAAAGGGGGNACC
35 GGGAGAACAAGCGTTTGAAGCGCA

SEQ ID NO:156

>GL3_21A_4_M13F Direction: N/A

CCATTTCAATACTATGTGGTCGTTACAATGCTAGAGGTAGGCACAGGGGGCGCAATGCAAGGGAGGAAGGGCGT
40 AACAAATTCCTGGCGCACGCTAACTTTCCAGGTGGCCAAGGCCACTGGTTATCAACCAGTTATTCCACGTTTTATT
TCCCTGGGTTGGAATTTACTGGAAAAAGGGGGCCAAGGGGGTTAAATTGGAGTATTCACTTTTATACAGACG
AGGGAAACAGGCCTCAGGGAGATAAGCTTACTTGACCCAAGTCTCTCGTCCTTAGGTCCATATCCAGGGAAACCC
AAAGGGAATTTTCCAAAACAGG

SEQ ID NO:157

>GL3_21A_1_M13F Direction: N/A

CCATTTCAATACTATGTGGTCGTTACAATGCTAGAGGTAGGCACAGGGGGCGCAATGCAAGGGAGGAAGGGCGT
AACATCTGCCACCTACTTCCAGGTGCCAAGCACTGTTATCAACATTATTCCACTTTATTCCCTGGTGATTATGAA
AGGCAGGTATTGATATTCACACTTAACAGACGAGGAAACAGCCTCAGGGAGATCAAGCTTACTTGACCCAGTCTC
50 TCTCCTAGTTCCATATCAGAACCAAGATTCAAACAGTTTGGTTTAGAAAATCTAGGATTTTTCAGCCATACCAA
AATAAAGTAGCCTCAGGGAATCAAAACATTACGAAAGAAGGTGACTTCTGAACTGAGTCTTGCAAGAATTCTGG
AAATTAACCTGGCAGTTGTATCAGGGGTGGTAAGGGGACGGAA

SEQ ID NO:158

>GL3_21A_2_M13R Direction: anti-sense

CCACTTGGATGAAAACCCCTTCAAGGATTACTGGATAGAATTCAGGCTTTCCCCTCTAGCCCCCAATCATAGCTCA
CAAACCTTCCTTGCTATTTGTTCTTAAGTAAAAAATCATTTTTCCTCCTCCCTCCCCAAACCCCAAGGAACTCTC
5 ACTCTTGCTCAAGCTGTTCCGTCCCCCTTACCACCCCTGATACAACAGGCTTAATTTCCAGAATTCTTGCAAG
ACTCAGTTTCAAGTACCTTCTTTCGGTGAATGTTTGGATTCCCTGAGGCTACTTTATTTTGGTATGGCTGAAA
AATCCTAGATTTTCTAAACAAAACCTGTTTGAATCTTGGTTCTGATATGGACTAGGAGAGAGACTG

SEQ ID NO:159

>GL3_21A_3_M13F Direction: N/A

CCATTTCAATACTATGTGGTTCGTTACAATGCTAGAGGTAGGCACAGGGGGCGCAATGCAAGGGAGGAAGGGCGTT
AACATCTGCCACCTACTTCCAGGTGCCAAGCACTGTTATCAACATTATTCCACTTTATTCCCTGGTGATTATGAA
AGGCAGGTATTGATATTTCACTTAACAGACGAGGAAACAGCCTTCAGGGAGATAAGCTTACTTGACCCAGTCT
15 CTCTTCCTAGTCCATATCAGAACCAAGATTCAAACAGGTTTGGTTTAGAAAATCTA

SEQ ID NO:160

>GL3_21A_3_M13R Direction: anti-sense

CCACTTGGATGAAAACCCCTTCAAGGATTACTGGATAGAATTCAGGCTTTCCCCTCTAGCCCCCAATCATAGCTCA
CAAAGCCTTCCTTGCTATTTGTTCTTAAGTAAAAAATCATTTTTCCTCCTCCCTCCCCAAACCCCAAGGAACTCT
20 CACTCTTGCTCAAGCTGTTCCGTCCCCCTTACCACCCCTGATACAACAGGCTTAATTTCCAGAATTCTTGCAA
GACTCAGTTTCAAGTACCTTCTTTCGTGAATGTTTGGATTCCCTGAGGCTACTTTATTTTGGTATGGCTGAAA
AATCCTAGATTTTCTAAACAAAACCTGTTTGAATCTTGGTTCTGATATGGACTAGGAGAGA

SEQ ID NO:161

>GL3_21A_4_M13R Direction: anti-sense

CCACTTGGATGAAAACCCCTTCAAGGATTACTGGATAGAATTCAGGCTTTCCCCTCTAGCCCCCAAGTCATAGCTC
ACAAAGCCTTCCTTGCTATTTGTTCTTAAGTAAAAAATCATTTTTCCTCCTCCCTCCCCAAACCCCAAGGGA
CTCCTCTTGCTCAAGGCTGTTCCGTTCGNCCTTACCACCCCTGATACAACAGGCTTAATTTCCAGAATTCT
TTGCAAGACTCAGTTTCAAGTACCTTCTTTCGTGAATGTTTGGATTCCCTGAGGCTACTTTATTTTGGTATGG
30 CTGAAAATCTAGATTTTCTAAACAAAACCTGTTTGAATCTTGGTTCTGATATGGACTAGGAGAGAGACTGGGT
CAAGTAAGCTTATCTCCCTGAGGCTGTTTCCT

SEQ ID NO:162

>GL3_21A_PCR_G3F1 Direction: N/A

CCATTTCAATACTATGTGGTTCGTTACAATGCTAGAGGTAGGCACAGGGGGCGCAATGCAAGGGAGGAAGGGCGTT
AACATCTGCCACCTACTTCCAGGTGCCAAGCACTGTTATCAACATTATTCCACTTTATTCCCTGGTGATTATGA
AAGGCAGGTATTGATATTTCACTTAACAGACGAGGAAACAGCCTCAGGGAGATAAGCTTACTTGACCCAGTCTC
TCTCCTAGTCCATATCAGAACCAAGATTCAAACAGGTTTGTGTTAGAAAATCTAGGATTTTTCAGCCATACCAAA
ATAAAGTAGCTCAGGGAATCAAACATTCACGAAAGAGGTGACTTCTGAACAGTCTTGGCAAGCAATCTGGA
40 AATTAACCTGGCAGTTGTATCAGGGGTGGTAAGGGGACGGAACAGCTTGAGCAAGAGTGAGAGTTTCTTGGGGT
TGGGGAGGGAGGAGGAAAAATGATTTTCTTAAAGAACAAATAGCAAGGAAGGTTTGTGAGCTAT

SEQ ID NO:163

>gi|5597012|ref|NM_000391.2| Homo sapiens ceroid-lipofuscinosis, neuronal
2, late infantile (Jansky-Bielschowsky disease) (CLN2), mRNA

ACATGACAGCAGATCCGCGGAAGGGCAGAATGGGACTCCAAGCCTGCCTCCTAGGGCTCTTTGCCCCCAT
CCTCTCTGGCAAATGCAGTTACAGCCCCGAGCCCGACCGAGCGGAGGACGCTGCCCCAGGCTGGGTGTCC
CTGGGCCGTGCGGACCCCTGAGGAAGAGCTGAGTCTCACCTTTGCCCTGAGACAGCAGAATGTGGAAGAC
TCTCGGAGCTGTTGTCAGGCTGTGTCGATCCCAGCTCTCCTCAATACGGAATAACCTGACCCCTAGAGAA
50 TGTGGCTGATCTGGTGAGGCCATCCCCACTGACCCCTCCACACGGTGCAAAAATGGCTCTTGGCAGCCGGA
GCCCAGAAGTGCCATTCTGTGATCACACAGGACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGC
TGCTGCTCCCTGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAACCCATGTTGTAAGGTCCCC
ACATCCCTACCAGCTTCCACAGGCTTGGCCCCCATGTGGACTTTGTGGGGGGACTGCACCGTTTTCCT
CCAACATCATCCCTGAGGCAACGTCCTGAGCCGAGGTGACAGGGACTGTAGGCCTGCATCTGGGGGTAA
55 CCCCCTCTGTGATCCGTAAGCGATACAACCTTGACCTCACAAGACGTTGGGCTCTGGCACCAGCAATAACAG
CCAAGCCTGTGCCAGTTTCTGGAGCAGTATTTCCATGACTCAGACCTGGCTCAGTTTATGCGCCTCTTC
GGTGGCAACTTTGCACATCAGGCATCAGTAGCCCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCCGGGA
TTGAGGCCAGCTAGATGTGAGTACCTGAGTGTGTTGGTCCCAACATCTCCACCTGGGTCTACAGTAG
CCCTGGCCGGCATGAGGGACAGGAGCCCTTCTGTCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTG
60 CCACATGTGCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA

ACACTGAGCTCATGAAGGCTGCCGCTCGGGGCTCTACCCCTGCTCTTCGCCTCAGGTGACAGTGGGGCCGG
GTGTTGGTCTGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAGCCCCATGTCAACACA
GTGGGAGGCACATCCTTCCAGGAACCTTTCTCATCACAAATGAAATTGTTGACTATATCAGTGGTGGTG
GCTTCAGCAATGTGTTCCACGGCCTTCATACCAGGAGGAAGCTGTAACGAAGTTCTTGAGCTCTAGCCC
5 CCACCTGCCACCATCCAGTTACTTCAATGCCAGTGGCCGTGCCTACCCAGATGTGGCTGCACTTTCTGAT
GGCTACTGGGTGGTCAGCAACAGAGTGGCCATTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGT
TTGGGGGGATCCTATCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCTCTTGGCTTTCTCAA
CCCAAGGCTCTACCAGCAGCATGGGGCAGGACTCTTTGATGTAACCCGTGGCTGCCATGAGTCTGTCTG
10 GATGAAGAGGTAGAGGGCCAGGGTTTCTGCTCTGGTCTGGCTGGGATCCTGTAACAGGCTGGGGAACAC
CCAACTTCCAGCTTTTGTGTAAGACTCTACTCAACCCCTGACCCTTTCTCTATCAGGAGAGATGGCTTGTG
CCTTGCCCTGAAGCTGGCAGTTTCACTCCCTTATTCTGCCCTGTTGGAAGCCCTGCTGAACCCCTCAACTAT
TGACTGCTGCAGACAGCTTATCTCCCTAACCTGAAATGCGGTGAGCTTGACTTGACTCCCAACCCCTACC
ATGCTCCATCATACTCAGGTCTCCCTACTCCTGCCTTAGATTCTCAATAAGATGCTGTGTAACCTATT
15 TTTGAATGCCTCTCCCTCCGCATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAG
ACTCTGTGCACTATTTCACTTGATATTCACTCCCAATTCACTGCAAGGAGACCTCTACTGTCACCGTTT
ACTCTTTCTTACCTTGACATCCAGAAACAATGGCCTCCAGTGCATACCTTCTCAATCTTTGCTTTATGGCC
TTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACCTCTCTGACTACTCTTGT
CTTCTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGCTCCATTTGTAGATTTTGTG
20 TCTTCTCAGTTTACTCATTTGTCCTTCTTCAAGCAATCACTGACATCTACAACCATTAACATCTCACTAAAT
AAGACTTTCTATCCAATAATGATTGATACCTCAAATGTAAGATGCGTGATACTCAACATTTTCATCGTCCA
CCTTCCCAACCCCAAAACAATTCATCTCGTTTCTTCTTGGTAAATGATGCTATGCTTTTCCAAACCAAGC
CAGAAACCTGTGTCATCTTTTACCCCCACCTTCAATCAACAAGTCCCTCAATCAACAAGTCCCTAGTACTG
CACATCTTAAATATATCTTTATCAGTCCACAAGTCCCTTCAATTATATTTCCCAAGTATATCTAGAACTT
25 ATCCACTTATATCCCCACTGCTACTACCTTAGTTTAGGGCTATATTCTCTTGAAAAAAGTGTCTTACT
TCCTGCCAATCCCCAAGTCATCTTCCAGAGTAAATGCAAATCCCATCAGGCCACTTGGATGAAAACCCCT
TCAAGGATTACTGGATAGAATTCAGGCTTTCCCTCCAGCCCCCAATCATAGCTCACAAACCTTCTCTTGC
TATTTGTTCTTAAGTAAAAAATCATTTTCTCTCTCCCTCCCCAAACCCCAAGGAACCTCTACTCTTGTCT
CAAGCTGTTCCGTCCCCTTACCACCCCTGATACAACCTGCCAGGTTAATTTCCAGAATTCTTGCAAGACTC
30 AGTTCAAGTCACTTCTTCTGTAATGTTTGTATCCCTGAGGCTACTTTATTTTGGTATGGCTGAAA
AATCCTAGATTTTCTAAACAAACCTGTTTGAATCTTGGTCTGATATGGACTAGGAGAGAGACTGGGTC
AAGTAAGCTTATCTCCCTGAGGCTGTTTCTCTGCTCTGTTAAGTGTGAATATCAATACCTGCCTTTTCA
TCACCAGGAATAAAGTGAATAATGTTGATAACAGTGCCTTGGCACCTGGAAGTAGGTGGCAGATGTTAA
CGCCCTTCTCTCCCTTGCAGTGCAGCCCCCTGTGCCTACCTCTAGCATTGTAACGACCACATAGTATTGAAA
35 TGGCCAGTTTACTTGTCTGCCTTCTTCCAAAGACCGTTGGTGCCTAGAGGACTAGAATCGTGTCTCTATT
TAACCTTGTGTTCCAGGTCTAGCTCAGGAGTTGGCAAATAAGAATTAAATGTCTGCTACACCGAAACA
AA

SEQ ID NO:164

>gi|5729770|ref|NP_000382.3| ceroid-lipofuscinosis, neuronal 2, late
40 infantile (Jansky-Bielschowsky disease) [Homo sapiens]
MGLQACLLGLFALILSGKCSYSPEPDQRRITLPPGWVSLGRADPEEELSLTFALRQONVERLSELVQAVSD
PSSPQYGYKLTLENVADLVRPSPLTLHTVQKWLAAAGAQKCHSVITQDFLTCWLSIRQAEILLPGAFFHH
YVGPTETHVVRSPHPYQLPQALAPHVDFVGGHLRFPTSSSLRQRPEPQVTGTVGLHLGVTPSVIRKRYN
45 LTSQDVSGTSSNSQACAQFLEQYFHDSDLAQFMRLFPGNFAHQASVARVVGQQGRGRAGIEASLDVQYL
MSAGANISTWVYSSPGRHEGQEPFLQWLMLLSNESALPHVHTVSYGDDDSLSSAYIQRVNTELMKAAAR
GLTLLFASGDSGAGCWSVSGRHQFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPS
YQEEAVTKFLSSSPHLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINE
50 HRILSGRPPLGLFNPRLYQHQHAGLFDVTRGCHESECLDEEVEGQGFCSGPGWDPVTGWGTPNFPALLKTL
LNP

SEQ ID NO:165

Novel protein disulfide isomerase
Novel (disulfide isomerase-like)
55 CGGACCAACACAGTATTGAGTCAACTGTGACCTTAAGATCAGAGGAACGTCAATACTGCCACAAGGCCACCTTTC
CAGAACTCGTGGGCAGGTAAACTATGCTTTGGATGTGCTTCTTTTACCAAATCACTCAACTCAGGAGCCACAA
ATAGTCCAGCAATTTTCACTTCCCTCAACGCTATTTTAGTCTCAAAGGAAACCATGTAAATTTTCATCAAGAGAAGG
TCAAAGGGGATATATCGCCACTGAAAATGTTTACACAGTGACCATGAGTTACACATTTACTTAGAGAACTTAAC
TTAATAAAGAATCTGTAGAGTGTGTTGGCTTGGAAAACACACACAAAGAAGATACCTCACGCTTAGTATGTTT
60 TGCTTTTCTGAACGCCACCCTGGGAACCCAGTGGCCTCTGTGGGACTGAACTCCTAAACGCAGGGTGCAGGAGC
TGGGCAGGAGAGGTGACCTCCAACCTGTGTTTCTAAAGTTCGTCTTTCGCTTGGCTCAGGACAAAGCGGTGTAACG

AGTCAAAGTCTCTGCCTCCACTGTGCTCACTGACTTTCTTCCCTCCTCGGAAAAGCAATAACGTGGGGTAGCCTC
GT

SEQ ID NO:166

5 Microtubule-associated protein 4 (MAP4)
>GL1_19_HIGH_3_G3F1 Direction: sense
CCACCTCCATTTCTGGAGGCAGACATACGTTCTTGATGAGAACTACTTCTGTTTCTGGAGGTGGAGTCACATCCT
TGCCCAGAGCCATTTCTGTTTCTGATAGTG

10 SEQ ID NO:167

>GL1_19_HIGH_3_G3R1 Direction: anti-sense
CACTATCAGAAACAGAAATGGCTCTGGGCAAGGATGTGACTCCACCTCCAGAAACAGAAGTAGTTCTCATCAAGA
ACGTATGTCTGCCTCCAGAAATGGAGGTGG

15 SEQ ID NO:168

>GL1_19_LOW_2_G3F1 Direction: sense
CCACCTCCATTTCTGGAGGCAGACATACGTTCTTGATGAGAACTACTTCTGTTTCTGGAGGTGGAGTCANATCCT
TGCCCAGAGCCATTTCTGTTTCTGATAGTG

20 SEQ ID NO:169

>gi|14195631|ref|NM_002375.2| Homo sapiens microtubule-associated protein 4
(MAP4), transcript variant 1, mRNA
TGCGACCGCCTCCCTGCGCCCCGCCCTCCGGCTAGCTCGCTGGCTCCCGGCTCCTCCCGACGTCTCCTA
CCTCCTCACGGCTCTTCCCGGCGCTCTCCTGGCTCCCTTCTGCCCCAGCTCCGTCTCGGCGGCGGGCGGGC
25 AGTTGCAGTGGTGCAGAAATGGCTGACCTCAGTCTTGAGATGCATTAACAGAACCATCTCCAGACATTGA
GGGAGAGATAAAGCGGGACTTCATTGCCACACTAGAGGCAGAGGCCCTTTGATGATGTTGTGGGAGAACT
GTTGGAGAAAAACAGACTATATTCCTCTCCTGGATGTTGATGAGAAAACCGGGAACCTCAGAGTCAAAGAAGA
AACCGTGCTCAGAAACTAGCCAGATTGAAGATACTCCATCTTCTAAACCAACACTCCTAGCCAATGTTGG
TCATGGAGTAGAAGGGAGCGATACTACAGGGTCTCCAATGAATTCCTTGAAGAGAAAAATGGCCTACCAG
30 GAATACCCAAATAGCCAGAACTGGCCAGAAGATACCAACTTTTGTTCCTAACCTGAGCAAGTGGTCGATC
CTATCCAGACTGATCCCTTTAAGATGTACCATGATGATGACCTGGCAGATTTGGTCTTTCCTCCAGTGC
GACAGCTGATACTTCAATATTTGTCAGGACAAAATGATCCCTTGAAAGACAGTTACGGTATGTCTCCCTGC
AACACAGCTGTTGTACCTCAGGGGTGGTCTGTGGAAGCCTTAACTCTCCACACTCAGAGTCTTTTGTTC
CCCCAGAGGCTGTTGCAGAACCTCCTCAGCCAACGGCAGTTCCCTTAGAGCTAGCCAAGGAGATAGAAAT
35 GGCATCAGAAGAGAGGCCACCAGCACAAGCATTGGAAATAATGATGGGACTGAAGACTACTGACATGGCA
CCATCTAAAGAAAACAGAGATGGCCCTCGCCAAGGACATGGCACTAGCTACAAAAACCGAGGTGGCATTGG
CTAAAGATATGGAATCACCCACCAAATTAGATGTGACACTGGCCAAGGACATGCAGCCATCCATGGAATC
AGATATGGCCCTAGTCAAGGACATGGAATACCCACAGAAAAAGAAGTGGCCCTGGTTAAGGATGTCAGA
TGGCCACAGAAACAGATGTATCTTCAGCCAAGAATGTGGTACTGCCACAGAAACAGAGGTAGCCCCAG
40 CCAAGGATGTGACACTGTTGAAAGAAACAGAGAGGGCATCTCCTATAAAAATGGACTTAGCCCCCTTCCAA
GGACATGGGACCACCCAAAGAAAAACAAGAAAGAAACAGAGAGGGCATCTCCTATAAAAATGGACTTGGCT
CCTTCCAAGGACATGGGACCACCCAAAGAAAAACAAGATAGTCCAGCCAAGGATTTGGTATTACTCTCAG
AAATAGAGGTGGCACAGGCTAATGACATTATATCATCCACAGAAATATCCTCTGCTGAGAAGGTGGCTTT
45 GTCTCAGAAACAGAGGTAGCCCTGGCCAGGGACATGACACTGCCCCCGGAAACCAACGTGATCTTGACC
AAGGATAAAGCACTACCTTTAGAAAGCAGAGGTGGCCCCCAGTCAAGGACATGGCTCAACTCCAGAAACAG
AAATAGCCCCCGCCAAGGATGTGGCTCCGTCCACAGTAAAGAAAGTGGGCTTGTGTAAGGACATGTCTCC
ACTATCAGAAACAGAAATGGCTCTGGGCAAGGATGTGACTCCACCTCCAGAAACAGAAAGTAGTTCTCATC
AAGAACGTATGTCTGCCTCCAGAAATGGAGGTGGCCCTGACTGAGGATCAGGTCCCAGCCCTCAAAACAG
AAGCACCCCTGGCTAAGGATGGGGTTCTGACCCCTGGCCAACAATGTGACTCCAGCCAAAGATGTTCCACC
50 ACTCTCAGAAACAGAGGCAACACCAAGTTCCAATTAAAGACATGGAAATTGCACAAACACAAAAAGGAATA
AGTGAGGATTCCCATTTAGAATCTCTGCAGGATGTGGGGCAGTCAGCTGCACCTACTTTTCATGATTTTAC
CAGAAACCATCACAGGAACGGGGAAAAAGTGCAGCTTGCCGGCCGAGGAGGATTCTGTGTTAGAAAAACT
AGGGGAAGGAAGGAACCATGCAACAGTCAACCTTCTGAGCTTTCTTCAGAGACCTCAGGAATAGCCAGGCCA
GAAGAAGGAAGGCCTGTGGTGGAGTGGGACAGAAATGACATCACCACCCACCGAACAAGGAGCTCCCAC
55 CAAGCCCAGAGAAGAAAAACAAGCCTTTGGCCACCCTCAACCTGCAAAGACTTCAACATCGAAAAGCCAA
AACACAGCCCACTTCTCTCCCTAAGCAGCCAGCTCCCACCACCATTTGGTGGGTGAATAAAAAACCCATG
AGCCTTGCTTCAGGCTTAGTGCCAGCTGCCCCACCAACGCCCTGCCGTGCGCTCTGCCAGGCCTTCCA
TCTTACCTTCAAAAGACGTGAAGCCAAAGCCCATTGCAGATGCAAAGGCTCCTGAGAAGCGGGCCTCACC
ATCCAAGCCAGCTTCTGCCCCAGCCTCCAGATCTGGGTCCAAGAGCACTCAGACTGTTGCAAAAACCCACA
60 ACAGCTGCTGCTGTTGCCTCAACTGGCCCAAGCAGTAGGAGCCCCCTCCACGCTCCTGCCCCAAGAAGCCCA

CTGCCATTAAAGACTGAGGGGAAAACCTGCAGAAGTCAAGAAGATGACTGCAAAGTCTGTACCAGCTGACTT
 GAGTCGCCCCAAGAGCACCTCCACCAGTTCCATGAAGAAAACCACACTCTCAGTGGGACAGCCCCCGCT
 GCAGGGGTGGTTCCCGCCGAGTCAAGGCCACACCCATGCCCTCCCGGCCCTCCACAACCTCTTCATAG
 ACAAGAAGCCCACCTCGGCCAAAACCCAGCTCCACCACCCCCGGCTCAGCCGCTGGCCACCAATACTTC
 5 TGCTCCTGATCTGAAGAATGTCCGCTCCAAGGTTGGCTCCACGGAAAACATCAAGCATCAGCCTGGAGGA
 GGCCGGGCCAAAGTAGAGAAAAAACAGAGGCAGCTGCTACAACCCGAAAGCCTGAATCTAATGCAGTCA
 CTAAAACAGCCGGCCCAATTGCAAGTGCACAGAAACAACCTGCGGGGAAAGTCCAGATAGTCTCCAAAAA
 AGTGAGCTACAGCCATATTGAGTCCAAGTGTGGTTCCAAGGACAATATTAAGCATGTCCCTGGAGGTGGT
 10 AATGTTTCAAGATTCAGAACAAAGAAAGTGGACATCTCTAAGGTCTCCTCCAAGTGTGGGTCTAAGGCTAACA
 TCAAGCACAAGCCTGGTGGAGGAGATGTCAAGATTGAAAGTCAGAAGTTGAACTTCAAGGAGAAGGCCCA
 GGCCAAGGTGGGATCCCTCGATAATGTGGGCCACCTACCTGCAGGAGGTGCTGTGAAGACTGAGGGCGGT
 GGCCAGCGAGGCTCCTCTGTGTCCGGTCCCCCTGCTGGGGAGGAGCCGGCCATCTCTGAGGCAGCGCCTG
 AAGCTGGCGCCCCCACTTCAGCCAGTGGCCTCAATGGCCACCCACCTGTCAAGGGGTGGTGACCAAG
 GGAGGCCAGACCTTGGACAGCCAGATCCAGGAGACAAGCATCTAATGATGACATTCTGGTCTCGTCTTC
 15 CGTCTCCCCCGTGTTCCTCTTGTCTCCCCGTGTTCCTCTCTCCCTTCCCTCCTCCCATGTCACTGCAGA
 TTGAGACCTACAGGCTGACGTTCGGGCAAATGCCAGGGCCCGCACCGACCACGGGGCCGACATTGTCTC
 CCGCCCCCACACTTCCCTGGCGGCCCAACTCGGGCTCCCGGTCTTGGCCCCCTTTCGGGGCTGTCTC
 CACTAGACCAGTGAGCGCTTGGGCGCCGTGTGGGCAGCCCGCTAGGCTCGCCTTCCCTCCTGCTTTGCG
 TGCCCCGGGGCAGCAGCAGCCCTGCCCCACACCTCCTCTCACTCCCCAGCCTGGGCCCCATCTCCCTGCTTT
 20 GGTCTTGGCCCCATCACTGCGCCACTGCTCCGTGGAGGAGGTGGGAGGGGGTGGGGTGGTTGAGGCTAA
 GTTGGGATCTAGGAGGAGGAGAACCAGATTCTATCCTCATCTTTTTTGGTTCTTTGGTCCAAACCCAAAA
 GAAACTGACATGCCCTCCCTTCTCCCTGGATCTACCTGGAGGGAAGAGTGGAGGTGGATTCCGAGTGGTG
 ACAGGACGCTGACCGTGGAGCTTAAGCCACTGCCTCTCCCTCTGGTCCCAAAATGGGCGCCCCCCCCCTC
 CCCATGCAGGTGGTGTGCGGCCCTTCTTGTCTGCCCTGCCCAAGTTGGGGGTGAGTGTGCTGCTGCCCA
 25 TGCTTAACATACCCGCTAGCTGCTGTACATTTTTCTTGTCTTTTGTCTTTTATTTTTTTCTAATAACCT
 AAAAAGTGGCAAAATAGTTCTGCAGGTTGAAGCCATGTCTACATGAAAGTCTCAGTAAGTGTAGAGGG
 AACAGGGCGGAGATATCCTTATGCCACCCCGCTGGAGGATGTGGGCAGCTTAGGGCCCTGGAGGCGGTG
 CGGCAGGGAAGAGGGGTGCAGAGGCTGTGGCTGGTGAGCCGTCAGGCACACAAGGGGGCCCTTGGAGCGT
 GGACTGGTTGGTTTTTGCATTTTTGTGTGTGTATGCTGCTTTTCTTTTCTAACAAGAGGCTGGTTTTGG
 30 CATCTGTCTCCATTCCCTGGGATCTGGTGGTCAGCCCTAGGATAAAAAGCCAGGGCTGGAGAACAAGAA
 AGGGCCAGGAGATGAATTC

SEQ ID NO:170

>gi|4505099|ref|NP_002366.1| microtubule-associated protein 4, isoform 1
 [Homo sapiens]
 MADLSLADALTEPSPDIEGEIKRDFIATLEAEAFDDVVGTVGKTDYIPLLDVDEKTGNSESKKKPCSET
 SQIEDTPSSKPTLLANGGHGVEGSDTTGSPTEFLEEKMAYQEYPNSQNWPEDTNFCFQPEQVVDPIQTDP
 FKMYHDDDLADLVFPSSATADTSIFAGQNDPLKDSYGMSPCNTAVVPQGSVEALNSPHSESFVSPEAVA
 EPPQPTAVPLELAKEIEMASEERPPAQALEIMMGLKTMDMAPSKETEMALAKDMALATKTEVALAKDMES
 40 PTKLDVTLAKDQPSMESDMALVKDMELPTEKEVALVKDVRWPTETDVSSAKNVVLPTETEVAAPAKDVT
 LKETERASPIKMDLAPSKDMGPPKENKKETERASPIKMDLAPSKDMGPPKENKIVPAKDLVLLSEIEVAQ
 ANDIISSTEISSAEKVALSSETEVALARDMTLPETNVILTCKALPLEAEVAPVKDMAQLPETEIAPAK
 DVAPSTVKEVGLLKDMSPLESETEMALGKDVTPPPETEVVLIKNVCLPPEMEVALTEDQVPALKTEAPLAK
 DGVLTLANNVTPAKDVPPLSETEATVPVPIKMEIAQTQKGISEDHLESQDVGQSAAPTFFMISPETITG
 45 TGKKCSLPAEEDSVLEKLGERKPCNSQPSSELSSETSGIARPEEGRPVVSGTGNIDITPPNKELPPSPEKK
 TKPLATTQPAKTSTSKAKTQPTSLPKQPAPTTIGGLNKKPMSLASGLVPAAPPKRPASARPSILPSKD
 VKPKPIADAKAPEKRASPSKPASAPASRSKSTQTVAKTTTAAAVASTGPSSRSPSTLLPKKPTAIKTE
 GKPAEVKMTAKSVPADLSRPKSTSTSSMKKTTTSLGTAPAAGVVPSPRVKATPMPSPSTTFIDKKPTS
 AKPSSTTPRLSRLATNTSAPDLKNVRSKVGSTENIKHQPGGGRKAVEKKTEAAATTRKPESNAVTKTAGP
 50 IASAQKQPAQKVQIVSKKVSYSHIQSKCGSKDNIKHVPGGGNVQIQNKKVDISKVSSKCGSKANIKHKPG
 GGDVKIESQKLNFKKAQAKVGS�DNVGHLPAGGAVKTEGGGSEAPLCPGPPAGEEPAISEAAPEAGAPT
 SASGLNGHPTLSGGGDQREAQTLDSQIQETS I

SEQ ID NO:171

>gi|14195625|ref|NM_030884.1| Homo sapiens microtubule-associated protein 4
 (MAP4), transcript variant 2, mRNA
 TGCGACCGCCTCCCTGCGCCCCGCCCCCTCCGGCTAGCTCGCTGGCTCCCGGCTCCTCCCGACGTCTCCTA
 CCTCCTCACGGCTCTTCCCGGCGCTCTCCTGGCTCCCTTCTGCCCCAGCTCCGTCTCGGCGGCGGGCGGC
 AGTTGCAGTGGTGAGAAATGGCTGACCTCAGTCTTGACATGATTAACAGAACCATCTCCAGACATTGA
 60 GGGAGAGATAAAGCGGGATTTTCATGGCCACTAGAGGCAGAGGCCTTTGATGATGTTGTGGGAGAACT
 GTTGGAAAAACAGACTATATTCCTCTCCTGGATGTTGATGAGAAAACCGGAACTCAGAGTCAAAGAAGA

AACCGTGCTCAGAACTAGCCAGATTGAAGATACTCCATCTTCTAAACCAACTCCTAGCCAATGGTGG
TCATGGAGTAGAAGGGAGCGATACTACAGGGTCTCCAAGTGAATTCCTTGAAGAGAAAATGGCCTACCAG
GAATACCCAAATAGCCAGAAGTGGCCAGAAGATACCAACTTTTGTTCCTAACCTGAGCAAGTGGTCCGATC
CTATCCAGACTGATCCCTTTAAGATGTACCATGTATGACCTGGCAGATTTGGTCTTTCCCTCCAGTGC
5 GACAGCTGATACTTCAATATTTGCAGGACAAAATGATCCCTTGAAAGACAGTTACGGTATGTCTCCCTGC
AACACAGCTGTTGTACCTCAGGGGTGGTCTGTGGAAGCCTTAAACTCTCCACACTCAGAGTCTTTGTTT
CCCCAGAGGCTGTTGCAGAACCTCTCAGCCAACGGCAGTTCCCTTAGAGCTAGCCAAGGAGATAGAAAT
GGCATCAGAAGAGAGGGCCACCAGCACAAGCATTGGAAATAATGATGGGACTGAAGACTACTGACATGGCA
10 CCATCTAAAGAAACAGAGATGGCCCTCGCCAAGGACATGGCCTAGCTACAAAACCGAGGTGGCATTTGG
CTAAAGATATGGAATCACCCACCAAATTAGATGTGACACTGGCCAAGGACATGCAGCCATCCATGGAATC
AGATATGGCCCTAGTCAAGGACATGGAATACCCACAGAAAAAGAAGTGGCCCTGGTTAAGGATGTCAGA
TGGCCACAGAAAACAGATGTATCTTCAGCCAAGAATGTGGTACTGCCACAGAAAACAGAGGTAGCCCCAG
CCAAAGGATGTGACACTGTTGAAAGAAAACAGAGAGGGCATCTCCTATAAAAATGGACTTAGCCCCCTCCAA
15 GGACATGGGACCACCCAAAGAAAACAAGAAAAGAAAACAGAGAGGGCATCTCCTATAAAAATGGACTTGGCT
CCTTCCAAGGACATGGGACCACCCAAAGAAAACAAGATAGTCCCAGCCAAGGATTTGGTATTACTCTCAG
AAATAGAGGTGGCACAGGCTAATGACATTATATCATCCACAGAAATATCCTCTGCTGAGAAGGTGGCTTT
GTCTCAGAAAACAGAGGTAGCCCTGGCCAGGGACATGACACTGCCCCCGGAAACCAACGTGATCTTGACC
AAGGATAAAGCACTACCTTTAGAAGCAGAGGTGGCCCCAGTCAAGGACATGGCTCAACTCCCAGAAAACAG
20 AAATAGCCCCCGCCAAGGATGTGGCTCCGTCCACAGTAAAGAAGTGGGCTTGTGTAAGGACATGTCTCC
ACTATCAGAAACAGAAATGGCTCTGGGCAAGGATGTGACTCCACCTCCAGAAACAGAAGTAGTTCTCATC
AAGAACGTATGTCTGCCTCCAGAAATGGAGGTGGCCCTGACTGAGGATCAGGTCCCAGCCCTCAAACAG
AAGCACCCCTGGCTAAGGATGGGGTTCTGACCTGGCCCAACAATGTGACTCCAGCCAAAGATGTTCCACC
ACTCTCAGAAACAGAGGCAACACCAGTTCCAATTAAGACATGGAAATTGCACAAACACAAAAAGGAATA
25 AGTGAGGATTTCCATTTAGAATCTCTGCAGGATGTGGGGCAGTCACTGTCACCTACTTTTCATGATTTTAC
CAGAAACCATCACAGGAACGGGAAAAAGTGCAGCTTGCCGGCCGAGGAGGATTCTGTGTTAGAAAACT
AGGGGAAAGGAAACCATGCAACAGTCAACCTTCTGAGCTTTCTTCAGAGACCTCAGGAATAGCCAGGCCA
GAAGAAGGAAGGCCTGTGGTGAGTGGGACAGGAAATGACATCACCACCCACCGAACAAGGAGCTCCCAC
CAAGCCCAGAGAAGAAAACAAGCCTTTGGCCACCACTCAACCTGCAAAGACTTCAACATCGAAAGCCAA
AACACAGCCCACCTTCTCTCCCTAAGCAGCCAGCTCCCACCACCATTTGGTGGGTGAATAAAAAACCCATG
30 AGCCTTGCTTCCAGCTTAGTGCCAGCTGCCCCACCCAAACGCCCTGCCGTGCGCTCTGCCAGGCCTTCCA
TCTTACCTTCAAAGAGCTGAAGCCAAAGCCCATTTGCAGATGCAAAGGCTCCTGAGAAGCGGGCCCTCACC
ATCCAAGCCAGCTTCTGCCCCAGCCTCCAGATCTGGGTCCAAGAGCACTCAGACTGTTGCAAAAACCACA
ACAGCTGCTGCTGTTGCCCTCAACTGGCCCAAGCAGTAGGAGCCCCCTCCACGCTCCTGCCCAAGAAGCCCA
35 CTGCCATTAAGACTGAGGGAACCTGCAGAAAGTCAAGAAGATGACTGCAAAGTCTGTACCAGCTGACTT
GAGTCGCCCCAAGAGCACCTCCACCAGTTCCATGAAGAAAACCACCACTCTCAGTGGGACAGCCCCCGCT
GCAGGGGTGGTTCCCAGCCGAGTCAAGGCCACACCCATGCCCTCCCGGCCCTCCACAACCTCTTTCATAG
ACAAGAAGCCCACCTCGGCCAAACCCAGCTCCACCACCCCCCGGCTCAGCCGCTGGCCACCAATACTTC
TGCTCCTGATCTGAAGAATGTCCGCTCCAAGGTGGCTCCACGGAAAAACATCAAGCATCAGCTTGGAGGA
GGCCGGTTTCAGATAGTCTCCAAAAAGTGAGCTACAGCCATATTCAGTCCAAGTGTGGTTCCAAGGACA
40 ATATTAAGCATGTCCCTGGAGGTGGTAATGTTTCAGATTTCAGAACAAAGAAAGTGGACATCTCTAAGTCTC
CTCCAAGTGTGGGTCTAAGGCTAACATCAAGCACAAGCCTGGTGGAGGAGATGTCAAGATTGAAAGTCAG
AAGTTGAACCTCAAGGAGAAGGCCAGGCCAAGGTGGGATCCCTCGATAATGTGGGCCACCTACCTGCAG
GAGGTGCTGTGAAGACTGAGGGCGGTGGCAGCGAGGCTCCTCTGTGTCCGGGTCCCCCTGCTGGGGAGGA
GCCGGCCATCTCTGAGGCAGCGCCTGAAGCTGGCGCCCCCACTTCAGCCAGTGGCCTCAATGGCCACCCC
45 ACCCTGTCAGGGGGTGGTGACCAAGGGAGGCCAGACCTTGGACAGCCAGATCCAGGAGACAAGCATCT
AATGATGACATTCTGGTCTCGTCTTCCGTCTCCCCCGTGTTCCTCTTGTCTCCCCCTGTTCCCCCTCTCC
CTTCCCTCCTCCCATGTCTACTGCAGATTGAGACCTACAGGCTGACGTTCCGGGCAAATGCCAGGGCCCCG
ACCGACCACGGGGCCGACATTGTCTCCCGCCCCCACACTTCCCTGGCGGCCCAACTCGGGCTCCCGGG
TCCTTGGCCCCCTTTCCCGGGCTGTCCACTAGACCAGTGAGCGCTTGGGCGCCGTGCTGGGCAGCCCCGCT
50 AGGCTCGCCTTCCCTCTGCTTTGCGTGCCCGGGGCGAGCAGCCCTGCCCCACACCTCCTCTCACTCC
CCAGCCTGGGCCCCTCTCCCTGCTTTGGTCTTGCCCCATCACTGCGCCACTGCTCCGTGGAGGAGGTGG
GAGGGGGTTGGGGTGGTTGAGGCTAAGTTGGGATCTAGGAGAGGAGAACCAGATTCTATCCTCATCTTTT
TTTGGTTCTTTGGTCCAAACCCAAAAGAACTGACATGCCCTCCCTTCTCCCTGGATCTACCTGGAGGGA
AGAGTGGAGGTGGATTCCGAGTGGTGACAGGACGCTGACCGTGGAGCTTAAGCCACTGCCTCTCCCTCTG
55 GTCCCACAAATGGGCGCCCCCTCCCATGCAAGTGGTGTGCGGCCCTTCTTGTGCTGCCCTGCCCAAG
TTGGGGTTCAGTGTGCTGCTGCTGCTTAACTAACATACCCGCTAGCTGCTGTACATTTTCTTGTGTTT
GTCTTTTATTTTCTTAATAACCTAAAACTGGCAAAATAGTTCTGCAGGTTGAAGCCATGCTCAT
GAAAGTCTCAGTAAGTGTAGAGGGAACAGGGCGGAGATATCCTTATGCCACCCCCGCTGGAGGATGTG
GGCAGCTTAGGGCCCTGGAGGCGGTGCGGCAGGGAAAGAGGGGTGCAGAGGCTGTGGCTGGTGAGCCGGTC
60 AGGCACACAAGGGGCCCTTGGAGCGTGGACTGGTGGTGTGTCATTTTGTGTTGTGTATGCTGCTTTTC
TTTCTAACCAAGAGGCTGGTTTTGGCATCTCTGTCCCATTCCTGGGATCTGGTGGTCAGCCCTAGGAT

AAAAAGCCAGGGCTGGAGAACAAGAAAGGCCAGGAGATGAATTC

SEQ ID NO:172

>gi|14195626|ref|NP_112146.1| microtubule-associated protein 4, isoform 2
[Homo sapiens]

MADLSLADALTEPSPDIEGEIKRDFIATLEAEAFDDVVGETVGKTDYIPLLDVDEKTGNSESKKKPCSET
SQIEDTPSSKPTLLANGGHGVEGSDTTGSPTEFLEEKMAYQEYPNSQNWPEDTNFCFQPEQVVDPIQTDP
FKMYHDDDLADLVFPSSATADTSIFAGQNDPLKDSYGMSPCNTAVVPQGSVEALNSPHSESFVSPEAVA
EPPQPTAVPLELAKBIEMASEERPPAQALEIMMGLKTTDMAPSKETEMALAKDMALATKTEVALAKDMES
PTKLDVTLAKDMQPSMESDMALVKDMELPTEKEVALVKDVRWPTETDVSSAKNVVLPTEVEVAPAKDVTL
LKETERASPIKMDLAPSKDMGPPKENKKETERASPIKMDLAPSKDMGPPKENKIVPAKDLVLLSEIEVAQ
ANDIISSTEISSAEKVALSSETEVALARDMTLPPETNVILTKDKALPLEAEVAPVKDMAQLPETEIAPAK
DVAPSTVKEVGLLKDMSPLESETEMALGKDVTPPPETEVVLIKVNCLPPEMEVALTEDQVPALKTEAPLAK
DGVLTLANNVTPAKDVPPLSETEATPVPIKDMEDIAQTQKGISEDHLESLODVGQSAAPTTFMISPETITG
TGKKCSLPAEEDSVLEKLGKPKCNSQPSSELSSETSGIARPEEGRPVVSGTGNDITTPPNKELPPSPEKK
TKPLATTQPAKTSTSKAKTQPTSLPKQPAPTTIGGLNKKPMSLASGLVPAAPPKRPASARPSILPSKD
VKPKPIADAKAPEKRASPSKPASAPASRSGSKSTQTVAKTTTAAAVASTGPSSRSPSTLLPKKPTAIKTE
GKPAEVKKMTAKSVPADLSRPKSTSTSSMKKTTTSGTAPAAGVVP SRVKATPMP SRPSTTPFIDKKPTS
AKPSSTTPRLSRLATNTSAPDLKNVRSKVGSTENIKHQPGGGRVQIVSKKVSYSHIQSKCGSKDNIKHVP
GGGNVQIQNKKVDISKVSSKCGSKANIKHKPGGVDVKIESQKLNFKKAQAKVGSGLDNVGHLPAGGAVKT
EGGGSEAPLCPGPPAGEEPAISEAAPEAGAPTSASGLNGHPTLSGGGDQREAOQLDSQIQETSI

SEQ ID NO:173

>gi|14195627|ref|NM_030885.1| Homo sapiens microtubule-associated protein 4
(MAP4), transcript variant 3, mRNA

TGCGACCGCCTCCCTGCGCCCCGCCCTCCGGCTAGCTCGCTGGCTCCCGGCTCCTCCCGACGTCTCCTA
CCTCCTCACGGCTCTTCCCGCGCTCTCCTGGCTCCCTTCTGCCCCAGCTCCGTCTCGGCGGCGGCGGGC
AGTTGCAGTGGTGAGAATGGCTGACCTCAGTCTTGAGATGCATTAACAGAACCATCTCCAGACATTGA
GGGAGAGATAAAGCGGGACTTCATTGCCACACTAGAGGCAGAGGCCTTTGATGATGTTGTGGGAGAAACT
GTTGGAAAAACAGACTATATTCCTCTCCTGGATGTTGATGAGAAAACCGGGAACCTCAGAGTCAAAGAAGA
AACCGTGCTCAGAAACTAGCCAGATTGAAGATACTCCATCTTCTAAACCAACACTCCTAGCCAATGGTGG
TCATGGAGTACAAGGGAGCGATACTACAGGGTCTCCAACCTGAATTCCTTGAAGAGAAAAATGGCCTACCAG
GAATACCCAAATAGCCAGAACTGGCCAGAGATACCAACTTTTGTGTTTCCAACCTGAGCAAGTGGTCCGATC
CTATCCAGACTGATCCCTTTAAGATGTACCATGATGATGACCTGGCAGATTTGGTCTTTCCCTCCAGTGC
GACAGCTGATACTTCAATATTTGCAGGACAAAATGATCCCTTGAAAGACAGTTACGGTATGTCTCCCTGC
AACACAGCTGTTGTACCTCAGGGGTGGTCTGTGGAAGCCTTAACTCTCCACACTCAGAGTCCTTTGTTT
CCCCAGAGGCTGTTGCAGAACCTCCTCAGCCAACGGCAGTTCCCTTAGAGCTAGCCAAGGAGATAGAAAT
GGCATCAGAAGAGAGGCCACCAGCACAAAGCATTGGAAATAATGATGGGACTGAAGACTACTGACATGGCA
CCATCTAAAGAAACAGAGATGGCCCTCGCCAAGGACATGGCACTAGCTACAAAAACCGAGGTGGCATTGG
CTAAAGATATGGAATCACCCACCAAATTAGATGTGACACTGGCCAAGGACATGCAGCCATCCATGGAATC
AGATATGGCCCTAGTCAAGGACATGGAACCTACCACAGAAAAAGAGTGGCCCTGGTTAAGGATGTGAGA
TGGCCACAGAAACAGATATCTTTCAGCCAGAAATGTGGTACTGCCACAGAAACAGAGGTAGCCCCAG
CCAAGGATGTGACACTGTTGAAAGAAAACAGAGAGGGCATCTCTATAAAAAATGGACTTAGCCCCCTTCCAA
GGACATGGGACCACCCAAAGAAAACAAGAAAAGAACAGAGAGGGCATCTCTATAAAAAATGGACTTGGCT
CCTTCCAAGGACATGGGACCACCCAAAGAAAACAAGATAGTCCCAGCCAAGGATTTGGTATTACTCTCAG
AAATAGAGGTGGCACAGGCTAATGACATTATATCATCCACAGAAATATCCTCTGCTGAGAAGGTGGCTTT
GTCTCTCAGAAACAGAGGTAGCCCTGGCCAGGGACATGACACTGCCCCCGGAAACCAACGTGATCTTGACC
AAGGATAAAGCACTACCTTTAGAAGCAGAGGTGGCCCCAGTCAAGGACATGGCTCAACTCCCAGAAACAG
AAATAGCCCCCGCCAAGGATGTGGCTCCGTCCACAGTAAAGAAAGTGGGCTTGTGTAAGGACATGTCTCC
ACTATCAGAAACAGAAATGGCTCTGGGCAAGGATGTGACTCCACCTCCAGAAACAGAAAGTAGTTCTCATC
AAGAACGTATGTCTGCCTCCAGAAATGGAGGTGGCCCTGACTGAGGATCAGGTCCCAGCCCTCAAACAG
AAGCACCCCTGGCTAAGGATGGGGTTCTGACCTGGCCAACAATGTGACTCCAGCCAAAGATGTTCCACC
ACTCTCAGAAACAGAGGCAACACCAGTTCCAATTAAAGACATGGAAATTGCACAAACACAAAAAGGAATA
AGTGAGGATTCCCATTAGAAATCTCTGCAGGATGTGGGGCAGTCAGCTGCACCTACTTTCATGATTTTAC
CAGAAACCATCACAGGAACGGGGAAAAAGTGCAGCTTGCCGGCCGAGGAGGATTCTGTGTTAGAAAAACT
AGGGGAAAGGAAACCATGCAACAGTCAACCTTCTGAGCTTTCTTCAGAGACCTCAGGAATAGCCAGGCCA
GAAGAAGGAAGGCCTGTGGTGAAGTGGGACAGGAAATGACATCACCACCCACCGAACCAAGGAGCTCCAC
CAAGCCGACAGAAAGAAAACAAGCCTTTGGCCACCACCTCAACCTGCAAAGACTTCAACATCGAAAGCCAA
AACACAGCCCACTCTCTCCCTAAGCAGCCAGCTCCCAACACCATTGGTGGGTTGAATAAAAAACCCATG
AGCCTTGCTTCAAGCTTAGTGCCAGCTGCCCCACCCAAACGCCCTGCCGTGCCTCTGCCAGGCCCTTCCA
TCTTACCTTCAAAGACGTGAAGCCAAAGCCCATTGCAGATGCAAAGGCTCCTGAGAAAGCGGGCTCACC

AAEKIERTVQIEASTVEIEERGVLRLTVVDTPGYGDAINCRDCFKTIISYIDEQFERYLHDESGLNRRH
IIDNRVHCCFYFISPFHGLKPLDVAFMKAIHNKNVIVPIAKADTLTLKERERLKKRILDEIEEHNKI
YHLPDAESDEDEDKEQTRLLKASIPFSVVGSNQLIEAKGKKVRGRLYPWGVVEVENPEHNDFLKLRMTL
ITHMQDLQEVTDLHYENFRSERLKRGGKRVENEDMNKDQILLEKEAELRRMQEMIARMQAQMOMQMOGG
5 DGDGGALGHV

SEQ ID NO:185

Novel (breast cancer antigen; tyrosine/serine/threonine kinase)

>G3_3_29_PCR_G3F1 Direction: N/A

10 CAGGGGGTGCTGAAGGCCCTCGAACTACATCCACCACATGGGATATGTACACAGGAGTGTCAAAGCCAGCCACAT
CCTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTGCAGCAACCTCAGCATGATAAGCCATGGGCAGCG
GCAGCGAGTGGTCCACGATTTTCCCAAGTACAGTGTCAAGGNTTCTGCCGTGGCTCAGCCCCGAGGTCCTCCAGC
AGAATCTCCAGGGTTATGATGCCAAGTCTGACATCTACAGTGTGGGAATCACAGCCTGTGAAC'TGGCCAACGGCC
15 ATGTCCCCTTTAAGGATATGCCTGCCACCCAGATGCTGCTAGAGAACTGAACGGCACAGTGGCCTGCCTGTTGG
ATACCAGCACCATCCCTGCTGAGGAGCTGACCATGAGCCCTTCGCGCTCAGTGGCCAACCTCTGGCCTGAGTGACA
GCCTGACCACCAGCACCCC

SEQ ID NO:186

>GL3_1.36_A2_M13F Direction: N/A

20 CAGGGGGTGCTGAAGGCCCTCGACTACATCCACCACATGGGATATGTACACAGGAGTGTCAAAGCCAGCCACATC
CTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTGCAGCAACCTCAGCATGATAAGCCATGGGCAGCGG
CAGCGAGTGGTCCACGATTTTCCCAAGTACAGTGTCAAGGTTCTGCCGTGGCTCAGCCCCGAGGTCCTCCAGCAG
AATCTCCAGGGTTATGATGCCAAGTCTGACATCTACAGTGTGGGAATCACAGCCTGTGAAC'TGGCCAACGGCCAT
GTCCCCTTTAAGGATATGCCTGCCACCCAGATGCTGCTAGAGAACTGAACGGCACAGTGGCCTGCCTGTTGGAT
25 ACCAGCACCATCCCTGCTGAGGAGCTGACCATGAGCCCTTCGCGCTCAGTGGCCAACCTCTGGCCTGAGTGACAGC
CTGACC

SEQ ID NO:187

>GL3_1.36_A2_M13R Direction: anti-sense

30 CAGAGGCTCAGAACTCCCAATCGTCCACCTCCAGCTCTTCCAGGTTTGTACCAGGCCAAAGATTCCACTGTGGT
CCTGAGACTGGCTGCCCTCAAAATTGGTGTATGGGGGTGACAGGACGAAGCAATTTCGGGCAAAGCCTCTGAGGCAC
GTCGCTTGATCTGGGGGAGAAGAGAGAGGTGGGTGACAGATCCTGTTGCTCTGGGTCCCAGNACACCATGGGGCA
AGGAACCATGGCCTGNTGGCAGACGGGCTGTCCGAGCCAACCTCCATGAGAGGAAGGAGCAGTGTCTTTCAGGGGC
TTCGGAAGNNGGGTAGTTTCTCTCATTCAAAANGAGGGGAGAGAAANCTGGCATCCGGGTTCGCTGAAGGCACT
35 GCTCCACAAAGTGGTGGAAGTGGGGGGAGAAGGTTTCGGTGGTAGGGGTGGGAGGGCGAGTACCCTAGGNGGGCC
GGGGGGTGCTG

SEQ ID NO:188

>GL3_1.36_B2_M13F Direction: N/A

40 CAGGGGGTGCTGAAGGCCCTCGACTACATCCACCACATGGGATATGTACACAGGAGTGTCAAAGCCAGCCACATC
CTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTGCAGCAACCTCAGCATGATAAGCCATGGGCAGCGG
CAGCGAGTGGTCCACGATTTTCCCAAGTACAGTGTCAAGGTTCTGCCGTGGCTCAGCCCCGAGGTCCTCCAGCAG
AATCTCCAGGGTTATGATGCCAAGTCTGACATCTACAGTGTGGGAATCACAGCCTGTGAAC'TGGCCAACGGCCAT
GTCCCCTTTAAGGATATGCCTGCCACCCAGATGCTGCTAGAGAACTGAACGGCACAGTGGCCTGCCTGTTGGAT
45 ACCAGCACCATCCCTGCTGAGGAGCTGACCATGAGCCCTTCGCGCTCAGTGGCCAACCTCTGGCCTGAGTGNCAGC
CTGACCAACAG

SEQ ID NO:189

>GL3_1.36_B2_M13R Direction: anti-sense

50 CAGAGGCTCAGAACTCCCAATCGTCCACCTCCAGCTCTTCCAGGTTTGTACCAGGCCAAAGATTCCACTGTGGT
CCTGAGACTGGCTGCCCTCAAAATTGGTGTATGGGGGTGACAGGACGAAGCAATTTCGGGCAAAGCCTCTGAGGCAC
GTCGCTTGATCTGGGGGAGAAGAGAGAGGTGGGTGACAGATCCTGTTGCTCTGGGTCCCAGGACACCATGGGGCA
AGGAACCATGGCCTGGTGGCAGACGGGCTGTCCGAGCCAACCTCCATGAGAGGAAGGAGCAGTGTCTTTCAGGGGC
TTCGGAAGCGGGTAGTTTCTCTCATTCAAAAGGAGGGGAGAGAAAGCTG

SEQ ID NO:190

>GL3_1.36_B3_M13F Direction: anti-sense

60 CAGAGGCTCAGAACTCCCAATCGTCCACCTCCAGCTCTTCCAGGTTTGTACCAGGCCAAAGATTCCACTGTGGT
CCTGAGACTGGCTGCCCTCAAAATTGGTGTATGGGGGTGACAGGACGAAGCAATTTCGGGCAAAGCCTCTGAGGCAC
GTCGCTTGATCTGGGGGAGAAGAGAGAGGTGGGTGACAGATCCTGTTGCTCTGGGTCCCAGGACACCATGGGGCA

AGGAACCATGGCCTGGTGGCAGACGGGCTGTCGGAGCCAACTCCATGAGAGGAAGGAGCAGTGTCTTTTCAGGGGC
TTCCGGAAGCGGGGTAGTTTCTCTCATTCAAAGGAGGGGAGAGAAAGCTGGCATCCGGGTTCGCGTGAAGGCACT
GCTCCACAAAGTGGTGGAAAGTGGGGGAGAAGGTTTCGGTGGTAGGGGTGGGAGGGCGAGTCACCGTNGGAGGGCC
GGGGGGTGTGGTGGTCANGCTGTCACTCAGGCCAGAGTTGGNCACTGATCGCGAANGGCTCATGGTCAGNTCCT
5 CANCAGGGATGGTGTGCTGGTATCCAACAGGCAGGGCACTGTGCCGTTTCTCTNAGCAGCATCTGGGTGGCA
GGCATATCCTTAAAGGGGACATGGCCGTTGGCCACTCACAGGCCTGT

SEQ ID NO:191

>GL3_1.36_B3_M13R Direction: sense

10 CAGGGGGTGTGAAGGCCCTCGACTACATCCACCACATGGGATATGTACACAGGAGTGTCAAAGCCAGCCACATC
CTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTGCGCAGCAACCTCAGCATGATAAGCCATGGGCAGCGG
CAGCGAGTGGTCCACGATTTTCCCAAGTACAGTGTCAAGGTTCTGCCGTGGCTCAGCCCCGAGGTCTTCCAGCAG
AATCTCCAGGGTTATGATGCCAAGTCTGACATCTACAGTGTGGGAATCACAGCCTGTGAACCTGGCCAACGGCCAT
GTCCCCCTTAAGGATATGCCTGCCACCCAGATGCTGCTAGAGAAACTGAACGGCACAGTGGCCTGCCTGTTGGAT
15 ACCAGCACCATCCCTGCTGAGGAGCTGACCATGAGCCCTTCGCGCTCAGTGGCCAACCTCTGGCCTGAGTGACAGC
CTGACCACCAG

SEQ ID NO:192

>GL3_1.36_B4_M13F Direction: anti-sense

20 CAGAGGCTCAGAACTCCCAATCGTCCACCTCCAGCTCTTCCAGGTTTGTACCAGGCCAAAGATTCCACTGTGGT
CCTGAGACTGGCTGCCCTCAAATTTGGTGTATGGGGGTGACAGGACGAAGCAATTCGGGCAAAGCCTCTGAGGCAC
GTCGCTTGATCTGGGGGAGAAGAGAGAGGTGGGTGACAGATCCTGTNGCTCTGGGTCCCAGGACACCATGGGGCA
AGGAACCATGGCCTGGTGGCAGANNAGTGTATAGCCCAACNCCATGAGAGGAAGGAGCAGTGTCTTTTCAGGGGC
TTCGGAAGCGGGTAGTTTCTCTCATTCAAAGGAGGGGAGAGAAAGCTGGCATCCGGGTTCGCGTGAAGGCACT
25 GCTCCACAAAGTGGTGGAAAGTGGGGGAGAAGGTTTCGGTGGTAGGGGTGGGAGGGCGAGTCACCGTTGGAGGGCC
GGGGGGTGTGGTGGTCAGGCTGTCACTCANGCCAGAGTTGGCCACTGACCGCGAAGGGCTCATGGTCAGTTCCT
CANCAGGGATGGTGTGTA

SEQ ID NO:193

>GL3_1.36_B4_M13R Direction: sense

30 CAGGGGGTGTGAAGGCCCTCGACTACATCCACCACATGGGATATGTACACAGGAGTGTCAAAGCCAGCCACATC
CTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTGCGCAGCAACCTCAGCATGATAAGCCATGGGCAGCGG
CAGCGAGTGGTCCACGATTTTCCCAAGTACAGTGTCAAGGTTCTGCCGTGGCTCAGCCCCGAGGTCTTCCAGCAG
AATCTCCAGGGTTATGATGCCAAGTCTGACATCTACAGTGTGGGAATCACAGCCTGTGAACCTGGCCAACGGCCAT
35 GTCCCCCTTAAGGATATGCCTGCCACCCAGATGCTGCTAGAGAAACTGAACGGCACAGTGGCCTGCCTGTTGGAT
ACCAGCACCATCCCTGCTGAGGAGCTGACCATGAGCCCTTCGCGCTCAGTGGCCAACCTCTGGCCTGAGTGACAGC
CTGACCACCAGCACCCCC

SEQ ID NO:194

>GL3_10.04_PCR_G3F1 Direction: N/A

40 CAGGGGGTGTGAAGGCCCTCGACTACGTCCACCACATGGGATATGTACACAGGAGTGTCAAAGCCAGCCACATC
CTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTGCGCAGCAACCTCAGCATGATAAGCCATGGGCAGCGG
CAGCGAGTGGTCCACGATTTNCCCAAGTACAGTGTCAAGGTTCTGCCGTGGCTCAGCCCCGAGGTCTTCCAGCAG
AATCTCCAGGGTTATGATGCCAAGTCTGACATCTACAGTGTGGGAATCACAGCCTGTGAACCTGGCCAACGGNCAT
45 GTCCCCCTTAAGGATATGCCTGCCACCCATATGCTGCTAGAGAAACTGAACGGCACAGTGGCCTG

SEQ ID NO:195

>GL3_3.29_A_M13F Direction: N/A

50 CAGGGGGTGTGAAGGCCCTCGACTACATCCACCACATGGGATATGTACACAGGAGTGTCAAAGCCAGCCACATC
CTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTGCGCAGCAACCTCAGCATGATAAGCCATGGGCAGCGG
CAGCGAGTGGTCCACGATTTTCCCAAGTACAGTGTCAAGGTTCTGCCGTGGCTCAGCCCCGAGGTCTTCCAGCAG
AATCTCCAGGGTTATGATGCCAAGTCTGACATCTACAGTGTGGGAATCACAGCCTGTGAACCTGGCCAACGGCCAT
GTCCCCCTTAAGGATATGCCTGCCACCCAGATGCTGCTAGAGAAACTGAACGGCACAGTGGCCTGCCTGTTGGAT
55 ACCAGCACCATCCCTGCTGAGGAGCTGACCATGAGCCCTTCGCGCTCAGTGGCCAACCTCTGGCCTGAGTGACAGC
CTGACCACCAGCACCC

SEQ ID NO:196

>GL3_3.29_A_M13R Direction: anti-sense

60 CAGAGGCTCAGAACTCCCAATCGTCCACCTCCAGCTCTTCCAGGTTTGTACCAGGCCAAAGATTCCACTGTGGT
CCTGAGACTGGCTGCCCTCAAATTTGGTGTATGGGGGTGACAGGACGAAGCAATTCGGGCAAAGCCTCTGAGGCAC

GTTCGCTTGATCTGGGGGAGAAGAGAGAGGTGGGTGACAGATCCTGTTGCTCTGGGTCCCAGGACACCATGGGGCA
AGGAACCATGGCCTGGTGGCAGACGGGCTGTGCGAGCCAACCTCCATGAGAGGAAGGAGCAGTGTCTTTCAGGGGC
TTCGGAAGCGGGGTAGTTTCTCTCATTCAAAAGGAGGGGAGAGAAAGCTGGCATCCGGGTTGCGCTGAAGGCACT
GCTCCACAAAGTGGTGGAAAGTGGGGGGAGAAGGTTTCGGTGGTAGGGGTGGGAGGGCGAGTCAACGTTGGAGGGCC
5 GGGGGGTGCTGGTGGTCAGGCTGTCACTCAGGCCAGAGTTGGCCACTGAGCGCGAAGGGCTCATGGTCAGCTCCT
CAGCAGGGATGGTGTGTTATCCAACAGGCAGGGCACTGTGCCGTTTCAGTTTCTCTAGCAGCATCTGGGTGGCAG
GCATATCCTTAAAGGGGACATGGCCGTTGGCCAGTTCACAGGCTGTGATTCCACACT

SEQ ID NO:197

>GL3_3.29_B_M13F Direction: N/A

CAGGGGGTGCTGAAGGCCCTCGACTACATCCACCACATGGGATATGTACACAGGAGTGTCAAAGCCAGCCACATC
CTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTGCAGCAACCTCAGCATGATAAGCCATGGGCAGCGG
CAGCGAGTGGTCCACGATTTTCCCAAGTACAGTGTCAAGGTTCTGCCGTGGCTCAGCCCCGAGGTCTCCAGCAG
AATCTCCAGGGTTATGATGCCAAGTCTGACATCTACAGTGTGGGAATCACAGCCTGTGAACTGGCCAACGGCCAT
15 GTCCCCCTTTAAGGATATGCCTGCCGCCAGATGCTGCTAGAGAACTGAACGGCACAGTGCCTGCCTGTTGGAT
ACCAGCACCATCCCTGCTGAGGAGCTGACCATGAGCCCTTCGCGCTCAGTGGCCAACCTCTGGCCTGAGTGACAGC
CTGACCACCAGCACCC

SEQ ID NO:198

>GL3_3.29_B_M13R Direction: anti-sense

CAGAGGCTCAGAACTCCCAATCGTCCACCTCCAGCTCTTCCAGGTTTGTACCAGGCCAAAGATTCCACTGTGGT
CCTGAGACTGGCTGCCCTCAAATTTGGTGTATGGGGGTGACAGGACGAAGCAATTCGGGCAAAGCCTCTGAGGCAC
GTCGCTTGATCTGGGGGAGAAGAGAGAGGTGGGTGACAGATCCTGTTGCTCTGGGTCCCAGGACACCATGGGGCA
AGGAACCATGGCCTGGTGGCAGACGGGCTGTGCGAGCCAACCTCCATGAGAGGAAGGAGCAGTGTCTTTTCAGGGGC
25 TTCGGAAGCGGGGTAGTTTCTCTCATTCAAAAGGAGGGGAGAGAAAGCTGGCATCCGGGTTGCGCTGAAGGCACT
GCTCCACAAAGTGGTGGAAAGTGGGGGGAGAAGGTTCCGTTGGTA

SEQ ID NO:199

>GL3_3.29_C_M13F Direction: N/A

CAGGGGGTGCTGAAGGCCCTCGACTACATCCACCACATGGGATATGTACACAGGAGTGTCAAAGCCAGCCACATC
CTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTGCAGCAACCTCAGCATGATAAGCCATGGGCAGCGG
CAGCGAGTGGTCCACGATTTTCCCAAGTACAGTGTCAAGGTTCTGCCGTGGCTCAGCCCCGAGGTCTCCAGCAG
AATCTCCAGGGTTATGATGCCAAGTCTGACATCTACAGTGTGGGAATCACAGCCTGTGAACTGGCCAACGGCCAT
GTCCCCCTTTAAGGATATGCCTGCCACCCAGATGCTGCTAGAGAACTGAACGGCACAGTGCCTGTTGGAT
35 ACCAGCACCATCCCTGCTGAGGAGCTGACCATGAGCCCTTCGCGCTCAGTGGCCAACCTCTGGCCTGAGTGACAGC
CTGACCACCAGCACCC

SEQ ID NO:200

>GL3_3.29_C_M13R Direction: anti-sense

CAGAGGCTCAGAACTCCCAATCGTCCACCTCCAGCTCTTCCAGGTTTGTACCAGGCCAAAGATTCCACTGTGGT
CCTGAGACTGGCTGCCCTCAAATTTGGTGTATGGGGGTGACAGGACGAAGCAATTCGGGCAAAGCCTCTGAGGCAC
GTCGCTTGATCTGGGGGAGAAGAGAGAGGTGGGTGACAGATCCTGTTGCTCTGGGTCCCAGGACACCATGGGGCA
AGGAACCATGGCCTGGTGGCAGACGGGCTGTGCGAGCCAACCTCCATGAGAGGAAGGAGCAGTGTCTTTTCAGGGGC
45 TTCGGAAGCGGGGTAGTTTCTCTCATTCAAAAGGAGGGGAGAGAAAGCTGGCATCCGGGTTGCGCTGAAGGCACT
GCTCCACAAAGTGGTGGAAAGTGGGGGGAGAAGGTTCCGTTGGTAGGGGTGGGAGGGCGAGTCAACGTTGGAGGGCC
GGGGGGTGCTGGTGGTCAGGCTGTCACTCAGGCCAGAGTTGGCCACTGAGCGCGAAGGGCTCATGGTCAGCTCCT
CAGCAGGGATGGTGTGTTATCCAACAGGCAGGGCACTGTGCCGTTCA

SEQ ID NO:201

>GL3_3.29_D_M13F Direction: N/A

CAGGGGGTGCTGAAGGCCCTCGACTACATCCACCACATGGGATATGTACACAGGAGTGTCAAANCCAGNCACATC
CTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTGCAGCAACCTCAGCATGATAAGCCATGGGCAGCGG
CAGCNAAGTGGTCCANGANN

SEQ ID NO:202

>gi|20559344|ref|XM_044379.3| Homo sapiens similar to serologically defined breast cancer antigen NY-BR-96 (LOC92335), mRNA

GCGGCGCGGCAGTAAACTGAGGAGGCGGAGCCAAGACGGTCGGGGCTGCTTGCTAACTCCAGGAACAGG
TTTAAGTTTTTGAAGTGAAGTAGGCCTACACAGTAGGAACTCATGTCAATTTCTTGTAAGTAAACCAGAG
60 CGAATCAGGCGGTGGGTCTCGGAAAAGTTCATTGTTGAGGGCTTAAGAGATTGGAACATTTGGAGAGC

AGCCTCCGGGTGACACTCGGAGAAAAACCAATGATGCGAGCTCAGAGTCAATAGCATCCTTCTCTAAACA
GGAGGTCATGAGTAGCTTTCTGCCAGAGGGAGGGTGTACGAGCTGCTCACTGTGATAGGCAAAGGATTT
GAGGACCTGATGACTGTGAATCTAGCAAGGTACAAACCAACAGGAGAGTACGTGACTGTACGGAGGATTA
ACCTAGAAGCTTGTTCATGAGATGGTAACATTCTTGCAGGGCGAGCTGCATGTCTCCAACTCTTCAA
5 CCATCCCAATATCGTGCCATATCGAGCCACTTTTATTGCAGACAATGAGCTGTGGGTGTGCATCATTC
ATGGCATAACGGTCTGCAAAAGATCTCATCTGTACACACTTCATGGATGGCATGAATGAGCTGGCGATTG
CTTACATCCTGCAGGGGGTGTGTAAGGCCCTCGACTACATCCACCACATGGGATATGTACACAGGAGTGT
CAAAGCCAGCCACATCCTGATCTCTGTGGATGGGAAGGTCTACCTGTCTGGTTTTCGCGAGCAACCTCAGC
10 ATGATAAGCCATGGGCAGCGGCAGAGTGGTCCACGATTTTCCCAAGTACAGTGTCAAGGTTCTGCCGT
GGCTCAGCCCCGAGGTCTCCAGCAGAATCTCCAGGGTTATGATGCCAAGTCTGACATCTACAGTGTGGG
AATCAGAGCCTGTGAATCGGCCACGGCCATGTCCCCTTTAAGGATATGCCCTGCCACCCAGATGCTGCTA
GAGAACTGAACGGCACAGTGGCCTGCTGTGGATACCAGCACCATCCCCGCTGAGGAGCTGACCATGA
GCCCTTCGCGCTCAGTGGCCAACCTTGGCCTGAGTGACAGCCTGACCACCAGCACCCCCCGGCCCTCCAA
15 CGGTGACTCGCCCTCCACCCCTACCACCGAACCTTCTCCCCCACTTCCACCCTTTGTGGAGCAGTGC
CTTCAGCGCAACCCGGATGCCAGGCCAGTGCAGCACCCCTCCTGAACCACTCTTTCTTCAAGCAGATCA
AGCGACGTGCCTCAGAGGCTTTGCCCGAATTGCTTTCGTCTGTACCCCCATCACCATTCTTGTAGGGCAG
CCAGTCTCAGGACCACAGTGAATCTTTGGCCTGGTAACAAACCTGGAAGAGCTGGAGGTGGACGATTGG
GAGTTCTGAGCCTCTGCAAACTGTGCGCATTTCTCCAGCCAGGGATGCAGAGGCCACCCAGAGGCCCTTCC
TGAGTGCCCGCCACATTTCCCGCCCTCCTGGGCAGATTGGGTAGAAAGGACATTCTTCCAGGAAAGTTGAC
20 TGCTGACTGATTGGGAAAGAAAATCCTGGAGAGACACTTCACTGCTCCAAGGCTTTTGAGACACAAGGGA
ATCTCAACAACAGGGATCAGGAGGGTCCAAAGCCGACATTCCAGTCTGTGAGCTCAGGTGACCTCCT
CCGAGAAGAGAGATGCTGCTCTGGCCCTGGGAGCTGAATTCCAAGCCAGGGTTTGGCTCCTTAAACCC
GAGGACCGCCACCTCTTCCAGTGCTTGCAGCCAGCCTCATTCTATTTAACTTTGCTCTCAGATGCCTCA
GATGCTATAGGTGAGTGAAGGGCAAGTAGTAAGTGCCTGCCTCCCTCCCTCAGACCTCTCCCTCATA
25 ATTCCAGAGAAGGGCATTCTGTCTTTTAAAGCACAGACTAAGGCTGGAACAGTCCATCCTTATCCCTCT
TCTGGCTTGGGCCCTGACACCTAAGTCTTTCCACGGTTTATGTGTGTGCCTCATTCTTTCCACCAAG
AATCCATCTTAGCGCCTCCTGCCAGCTGCCCTGGTGCTTCTCCAAGGGCCATCAGTGTCTTGCTAGCT
TGAGGGCTTAAGTCCTTATGCTGTGTAGTTTCGTTGTCAGAA

30 SEQ ID NO:203

>gi|14770985|ref|XP_044379.1| similar to serologically defined breast
cancer antigen NY-BR-96 [Homo sapiens]
MSFLVSKPERIRRWVSEKFIVEGLRDLLELFGEPQPDTRRKTNDASSESIASFQKQEVMSFLPEGGCYE
LLTVIGKGFEDLMTVNLARYKPTGEYVTVRRINLEACSNEVMVTFLOGLHVSCLFNHPNIVPYRATFIAD
35 NELWVVTSF MAYGSAKDLICTHFM DGMNELAIAYILQGV LKALDYIHHMGYVHRSVKASHILISVDGKVY
LSGLRSNLSMISHGQRQRVVHDFPKYSVKVLPWLSPEVLQQLQGYDAKS DIYSVGITACELANGHV PFK
DMPATQMLLEKLN GTVPCLLD TSTIPA EELTMSPSRSVANSGLSDSLTSTPRPSNGDSPSHPYHRTFSP
HFHHFVEQCLQRPNDARPSASTLLNHSFFKQIKRRASEALPELLRPVTPITNFEQSQSDHSGIFGLVTN
40 LEELEVDDWEF

40 SEQ ID NO:204

Novel (chromosome 1)

>GL2_46A_B12_G3F1 Direction: sense

ATCTACATGAGGTCTGTAGATTGAGCAACTTTGGCTATTTGGGTAGCTGATCCACTGCCAAGCTGGCTCCCAGC
45 TACCACCTGTTACTTTGACAACTCTAGCTTCAGGGCTACCAGGACCATTGCAAACTTACACCCACAGAAAACT
CAAAGTCTCCAGAGGCATAAGCTCCAATCCCCCAGAGAGCTCGCACTAGCATTCTTCTCCCTTTCTAGTAGTACT
TTTATGGGGGGGAGTTATTCTTCCCCACAAAAGTGAGCGGCGACAATAAAAACAATGTTTCCCACAAAGAGTTTT
CCCCGGTTCTTCTCCAAAAATTNNGGGGGGGCCCCCGTTTNTCCCCCGGGCGACGGGGGGCGGNGGGCCCCGCC
50 CCAAAACCACAANGNGGTGTGTTCCCTCTGAAAGCTCNNACAAAGAGGGGTGACCCCCCGGTTTGGGTGA
TGAATACCGCCCCGAAGATAGGNCGCCGAAAAAGAGCCAACCCNNGAAAAGAGAGATTTTATATAGAGGAAGCA
AAAACGAAANGANAGCGTTTTTTTTTTTTTGTCCNGCCAACAACATGGGGGGCCCCCGTTTTATAAATAGAGTT
TTTTGTTTTTAAACAATCGACTTTTTTTTTTCCCCCGTGATTTAAAAAAGAAGGAAGGGTATTTATATTNTT
GGGGGNGCGTGATTATATAAATTATTTAGTGGGGTGCGGGGANGGAAAGAGNGAAGCTGTTTATCCCATCAAAA
55 TATTATTGTG

55 SEQ ID NO:205

>GL2-30-CON-TOPO Direction: sense

ATCTACATGAGGTCTGTAGATTGAGCAACTTTGGCTATTTGGGTAGCTGATCCACTGCCAAGCTGGCTCCCAGC
60 TACCACCTGTTACTTTGACAACTCTAGCTTCAGGGCTACCAGGACCATTGCAAACTTACACCCACAGAAAACT
CAAAGTCTCCAGAGGCATAAGCTCCAACCCCCAGAGAGCTCCACAGCATTCTTCTCCTTCTAGTAGTACTTTTG

TTTCCAGGAAAATCATCTCATGCTCGGGCCACAGTCACAGTCCTGACCAGCACCATAGAAAATTTTCAGCCATTATT
CTAGTTGTATGGAGTCATTCAAATACTAGTTCCGAAAGAAGTTTCAAAGGTCCAACGCCGAGCATCTTTTACCT
TTGCTAATTTTTTCCATAAATTGAAAGCCTT

5 SEQ ID NO:206

>GL2_46A_B07_G3F1 Direction: N/A

ATCTACATGAGGTCCTGTAGATTGAGCAACTTTGGCTATTTGGGTAGCTGATCCACTGCCAAGCTGGCTCCCAGC
TACCACCTGTTACTTTGACAACTCTAGCTTCAGGGGGCTACCAGGACCATTGCAAAACTTACACCCACAGAAAA
CTCAAAGTCTCCAGAGGCATAAGCTCCAACCCCCAGAGAGCTCCACAGCATTCTTCTCCTTCTAGTAGTACTTTT
10 GGTTCAGGAAAATCATCTCATGCTCGGGCCACAGTCACAGTCCTGACCAGCACCATAGAAAATTTTCAGCCATTA
TTCTAGTTGTATGGAGTCATTCAAATACTAGTTCCGAAAGAAGTTTCAAAGGTCCAACGCCGAGCATCTTTT
ACCTTTGCTAATTTTTT

SEQ ID NO:207

15 >GL2_46A_B08_G3F1 Direction: sense

ATCTACATGAGGTCCTGTAGATTGAGCAACTTTGGCTATTTGGGTAGCTGATCCACTGCCAAGCTGGCTCCCAGC
TACCACCTGTTACTTTGACAACTCTAGCTTCAGGGGCTACCAGGACCATTGCAAAACTTACACCCACAGAAAACT
CAAAGTCTCCAGAGGCATAAGCTCCAACCCCCAGAGAGCTCCACAGCATTCTTCTCCTTCTAGTAGTACTTTTGG
20 TTCCAGGAAAATCATCTCATGCTCGGGCCACAGTCACAGTCCTGACCAGCACCATAGAAAATTTTCAGCCATT

SEQ ID NO:208

>GL2_46A_B10_G3F1 Direction: sense

ATCTACATGAGGTCCTGTAGATTGAGCAACTTTGGCTATTTGGGTAGCTGATCCACTGCCAAGCTGGCTCCCAGC
TACCACCTGTTACTTTGACAACTCTAGCTTCAGGGGCTACCAGGACCATTGCAAAACTTACACCCACAGAAAACT
25 TCAAAGTCTCCAGAGGCATAAGCTCCAACCCCCAGAGAGCTCCACAGCATTCTTCTCCTTCTAGTAGTACTTTTGG
GTTTCAGGAAAATCATCTCATGCTCGGGCCACAGTCACAGTCCTGACCAGCACCATAGAAAATTTTCAGCCATTAT
TCTAGTTGTATGGAGTCATTCAAATACTAGTTCCGAAAGAAGTTTCAAAGGTCCAACGCCGAGCATCTTTTACCT
TTTGCTAATTTTTTCCATAAATTGAAAGCCTT

30 SEQ ID NO:209

>GL2_46A_B11_G3F1 Direction: N/A

AATCTACATGAGGTCCTGTAGATTGAGCAACTTTGGCTATTTGGGTAGCTGATCCACTGCCAAGCTGGCTCCCAG
CTACCACCTGTTACTTTGACAACTCTAGCTTCAGGGCTACCAGGACCATTGCAAAACTTACACCCACAGAAAACTCA
AAGTCTCCAGAAGCATAAGCTCCAACCCCCAGANAGCTCCAAGATTCTTCTCCTTCTAGTAGTACTTTTGGTTTC
35 CAGGAAAATCATCTCATGCTCGGGCCACAGTCACAGTCCTGACCAGCACCATAGAAAATTTTCAGCATTTCTTAGT
TGTATGTGAGTCCTTCAAATACTAGTTCCGAAAGAAGTTTCAAAGGTCCAACGCCGAGCATCTTTTACNTTGGT
AATTTTTCCCTAAATTGAAAGCCTTCCGCAGAAACCAGCACAGTGGTTAGATAGATAAA

SEQ ID NO:210

40 Clone 30 -Novel (Chromosome 1)>

ATCTACATGAGGTCCTGTAGATTGAGCAACTTTGGCTATTTGGGTAGCTGATCCACTGCCAAGCTGGCTCCCAGC
TACCACCTGTTACTTTGACAACTCTAGCTTCAGGGGCTACCAGGACCATTGCAAAACTTACACCCACAGAAAACT
CAAAGTCTCCAGAGGCATAAGCTCCAACCCCCAGAGAGCTCCACAGCATTCTTCTCCTTCTAGTAGTACTTTTGG
TTTCCAGGAAAATCATCTCATGCTCGGGCCACAGTCACAGTCCTGACCAGCACCATAGAAAATTTTCAGCCATTATT
45 CTAGTTGTATGGAGTCATTCAAATACTAGTTCCGAAAGAAGTTTCAAAGGTCCAACGCCGAGCATCTTTTACCT
TTGCTAATTTTTTCCATAAATTGAAAGCCTT

SEQ ID NO:211

Novel (chromosome 3; h40)

50 >GL2_46B_C03_G3F1 Direction: N/A

GCGTCGCTCGGCGTTAGCCAAGGCCCGGGCGGCCACCCTCCGGGGGCACTAGGTCTGGGGCCGAGTGCCCCC
CCATAGGCACACAGAGNGAAAGCGCCAATGGGCGTGGGTTTTTTTTTAAATTTCCGGGGCGGGGCGNAACCCCGG
AACACAACGGGCAAACCTGTACAATTTTTGTGGGGANCAACAACCGGGACACAANTTTTTGGGGNAATGGCGCCA
CAAAAAAANAACCAACGACGCGACGAAAGCGCGTGGTGTGTCNGAATACTTTTTATAAAACCACCAATGG
55 GCGCCTCTCGCTTGTGCGCGCCCTTCTCCCAAANGGGGGCGGCGCTGTTTTTCCCCAACACGGGCGTGGGG
CCCCAAAAAATTTGTGGGCGCCCAAGAAAAAATATTTTAATAAAACAGAGGCGGCCCCCAACACAGGCGGT
TNTGGGGGAACAGGAAACCAACACAACGACAACACGGCGCTGGCGGCAACAAGCAAACACGGGAGCAAAACAAAC
ACGGGCGGAAAGAAACACACGGGCGAAACGACTTTTATAACAAACACTTTTGTGGGGCGCAAACAGCACACAC
ACAACCAAAATGGGCGCCCCCCCCNCCGAGGGGGCCCAAAAGGNCGAAGAGAAAAAACACCCCCCCCCAAAA
60 AGAGAGCGCCCGCAAAACAACCAAAAAAAGGGGAGGAGTTTGGAGAGCAGGGCGAGANGTTTTTTTATATTA

AAACACGGGCNACGAAAACATCATTTTATAAATAGGGCAGGCACAAAAACACGATTTATAATATTTCAACAGAAAA
CAGGGNTGGCCTCTGGGGGTTTGGTGTGTGTGCTTCCCNACCGGTGGTGCGCGTGTTCCTCCTGGTGNGGAA
AAAAGACACCTGTTTTATAAAAGAGGAGNGTATTNCCNCCTTTTTGTGGGTGCGAAAAACAGACNGAGGCGGGA
NGATNTATCCCACTCTTTTGGTGGTGCAACAAGTGTTTTATATAGNATACCTG

SEQ ID NO:212

>GL2_46B_C01_G3F1 Direction: N/A

GCGTTGCTCGGCGTTAGCCAAGGGCCCGGGCGGCGCCACCCTCCGGGGGCACTAGGTCTGGGGCCGCAGTGCCCA
GNCACAGAGCAGCGTTTATCTGGGACGAAGATGAATGGAAAGAATTGGAGCAAAAAGAGGTTGATTACAGCGGGC
TCANGGGTTCAGGCGAATGCAAATAAGCACGTGAAAAGGAAGAAGACGATAATGAAAAGCCGCAGAAACCAGCAC
AGTGTTAGAT

SEQ ID NO:213

>GL2_46B_C02_G3F1 Direction: sense

CAAGCCGGTGACGCGTTGTCTAGTGTCTGGGTCTGCGGGCGTTGCTTCGGCGTTAGCCAAGCCGGGCGGCGCCA
CCCTCCGGGGGGGGGGCCACTAGGTCTGGGGCCGCAGTGCCACAGACAGCGGTTTATCGGGACGAAGAAT
GAATTGGAAAGAATTGGAGCAAAAAGAGTTGATAACAGCGGCCTCAGGGGTTTCAGGCAATGCAAATAAGCAGTG
AAAAGGAAGAAGACGATAATGAAAAGCCGCAGAAACCAGCACAGTGGTTAGATAGATAAAGCGGGCGCTCGACTA
GTCTGAGGTCTGATA

SEQ ID NO:214

>GL2_46B_C04_G3F1 Direction: N/A

GCGTCTGCTCGGCGTTAGCCAATGCCCGGGCGGCGCCACCCTCCGGGGGCACTAGGTCTTGGGGCCGCAGTGCCACAG
ACAGAGCAGCGTTTATCGGGACGAAGATGAATGGAAAGAATTGGAGCAAAAAGAGGTTGATTACAGCTGGCCCTCA
GGGGGTTTCAGGCAATGCAAATAAGCAGTGAAAAGGAAGAAGAC

SEQ ID NO:215

>GL2_46B_C05_G3F1 Direction: sense

GCGTCTGCTCGGCGTTAGCCAAGGCCCGGGCGGCGCCACCCTCCGGGGGCACTAGGTCTGGGGCCGCAGTGCCACAG
CACAGAGCAGCGTTTATCGGGACGAAGATGAATGGAAAGAATTGGAGCAAAAAGAGGTTGATTACAGCGGCTCAG
GGGTTTCAGGCAATGCAAATAAGCAGTGAAAAGGAAGAAGACGATAATGAAAAG

SEQ ID NO:216

>GL2_46B_C06_G3F1 Direction: sense

GCGTCTGCTCGGCGTTAGCCAAGGCCCGGGCGGCGCCACCCTCCGGGGGCACTAGGTCTGGGGCCGCAGTGCCACAG
CACAGAGCAGCGTTTATCGGGACGAAGATGAATGGAAAGAATTGGAGCAAAAAGAGGTTGATTACAGCGGCTCA
GGGGTTTCAGGCAATGCAAATAAGCAGTGAAAAGGAAGAAGACGATAATGAAAAG

SEQ ID NO:217

>GL2_95_1_M13R Direction: anti-sense

CTTTTCATTATCGTCTTCTTCTTTTCACTGCTTATTTGCATTGCCTGAACCCTGAGGCCGCTGTAATCAACCTC
TTTTTGCTCCAATTCCTTCCATTTCCTTCTGCTCCGATAAACGCTGCTCTGTGCTGGGCACTGCGGCCCCAGACC
TAGTGCCCCCGGAGGTTGGCGCCGCGGCGCTCGGCTAACGCCGAGCGACGC

SEQ ID NO:218

Clone 95 -Novel (Chromosome 3/ H41)>

GCGTCTGCTCGGCGTTAGCCAAGGCCCGGGCGGCGCCACCCTCCGGGGGCACTAGGTCTGGGGCCGCAGTGCCACAG
CACAGAGCAGCGTTTATCGGGACGAAGATGAATGGAAAGAATTGGAGCAAAAAGAGGTTGATTACAGCGGCTCA
GGGGTTTCAGGCAATGCAAATAAGCAGTGAAAAGGAAGAAGACGATAATGAAAAG

SEQ ID NO:219

Novel (chromosome 17)

>GL2_147_1_M13F Direction: anti-sense

CTTGCCATTCCACTCCAGCCTGGGCAACGAGAGCGAAACTCTGTCTCAAAAAGAAAAGAAAAGAAAAGAAAACC
ACCTCCCACCAGGTTAAATAAATAATTAACATTTTGGTATGTACCCTTCCAGACATTTTCTATGCATCACATCA
ATAAATATATGATAGGATATTTTACATTTGATGTATCTGAAGATTAATTCAACAAATGTTTAGTGAGTAGGCTT
GTTGTAGGTGCTGGGAATTCAGCAGTGAACAAAACAAAGTCTCTGCCTCATGGAGCTCATATTATAGTAGGGAAG
ACTGTCAATAAGCAATATACACATTGACTAAATGATAGATAATGCCAGTAATGTGGCAAGTTTTATGGAAAAT
AAATCAGGATAAAAGGATAAGAGTGATATGGGGATTCTCTTAGTTAGGGTAGATAAGGAAGGCCCTCTTAGCTAAG

GAGGCACATAAATATCTGATCATCTAATGAGAGAAGGAGCAACATGGGATTCTAGAGGGAACTATCCTAGAAA
GA

SEQ ID NO:220

>GL2_147_2_M13F Direction: sense

CTGCTTCTCTTCTAGGATAGTTTCCCTCTAGAAATCCATGTTGCTCCTTCTCTCATTAAATGATCAGATATTTAT
GTGCCTCCTTAGCTAAGAGGCCTTCCTTATCTACCCTAACTAAGAGAATCCCCATATCACTCTTATCCTTTTATC
CTGATTTATTTTTCCATAAACTTGCCACATTACTGGGCATTATCTATCATTTAGTAATGTGTATATTGCTTATT
GACAGTCTTCCCTACTATAATATGAGCTCCATGAGGCAGAGACTTGTGTTTGTTCACCTGCTGAATTCCCAGCACC
TACAACAAGCCTACTCACTAAACATTTGTTGAATTAATCTTCAGGATACATCAAATGTAAAAATATCCTATCATA
TATTTATTGATGTGATGCATAGGAAAATGTCTGGAAGGGTACATACCAAAATGTTAATTATTTATTTAACCTGGT
GGGAGGTGGATTATTTCTTTCTTTCTTTCTTTTGGAGACCAGAGTTTCCGCATCTTCGTTGCCCCAGGCTGGG
AGTGGAATTTGGGCAAAGCCCGCAGAAACCCCAAGCAACAAGTGGGTTTAGATAGNATCAAAGCGGGCCGCTCGA
CTTAANTCCTTGAAGGTCTGATTACTCAACCTGACTTGTCAGAAA

SEQ ID NO:221

>GL2_147_2_M13R Direction: anti-sense

CTTGCCATTCCTACTCCAGCCTGGGCAACGAGAGCGAACTCTGTCTCAAAAAGAAAAGAAAAGGAAAGAAAAACC
ACCTCCCACCAGGTTAAATAAATAATTAACATTTTGGTATGTACCTTCCAGACATTTTCCTATGCATCACATCA
ATAAATATATGATAGGATATTTTACATTTGATGTATCCTGAAGATTAATTCAACAAATGTTTAGTGAGTAGGCTT
GTTGTAANGTGCTGGGAATTCAGCAGTGAACAAAACAAAGTCTCTGCCTCATGGAGCTCATATTATAGTAGGGAA
GACTGTCAATAAGCAATATACACATTACTAAATGATAGATAATGCCCAGTAATGTGGCAAGTTTATGGAAAAAT
TAATCAGGATAAAGGATAAGAGTGATATGGGGATTCTCTTAATTAGGGTAGATNAGGAAGGGCTCTTTAGCTNAA
GGAAGCAACATAAATATCTGATCATTTTAAAANGGGAAGGAACGAAAAAGGGGAAAGCCAAAACACAAATGGGGG
ACAATTTATACCTTACGAAAAGGGGGGAAAAAACCATAAAATCCCTAAGGAAACANAAAAGCAAAAGCNCAAAGA
CCAGCAAAAACAA

SEQ ID NO:222

>GL2_147_3_M13F Direction: anti-sense

CTTGCCATTCCTACTCCAGCCTGGGCAACGAGAGCGAACTCTGTCTCAAAAAGAAAAGAAAAGGAAAGAAAAACC
ACCTCCCACCAGGTTAAATAAATAATTAACATTTTGGTATGTACCTTCCAGACATTTTCCTATGCATCACATCA
ATAAATATATGATAGGATATTTTACATTTGATGTATCCTGAAGATTAATTCAACAAATGTTTAGTGAGTAGGCTT
GTTGTAGGTGCTGGGAATTCAGCAGTGAACAAAACAAAGTCTCTGCCTCATGGAGCTCATATTATAGTAGGGAAG
ACTGTCAATAAGCAATATACACATTACTAAATGATAGATAATGCCCAGTAATGTGGCAAGTTTATATGGAAAAAT
AAATCAGGATAAAGGATAAGAGTGATATGGGGATTCTCT

SEQ ID NO:223

>GL2_147_3_M13R Direction: sense

CTGCTTCTCTTCTAGGATAGTTTCCCTCTAGAAATCCATGTTGCTCCTTCTCTCATTAAATGATCAGATATTTAT
GTGCCTCCTTAGCTAAGAGGCCTTCCTTATCTACCCTAACTAAGAGAATCCCCATATCACTCTTATCCTTTTATC
CTGATTTATTTTTCCATAAACTTGCCACATTACTGGGCATTATCTATCATTTAGTAATGTGTATATTGCTTATT
GACAGTCTTCCCTACTATAATATGAGCTCCATGAGGCAGAGACTTGTGTTTGTTCACCTGCTGAATTCCCAGCACC
TACAACAAGCCTACTCACTAAACATTTGTTGAATTAATCTTCAGGATACATCAAATGTAAAAATATCCTATCATAT
ATTTATTGATGTGATGCATAGGAAAATGTCTGGAAGGGTACATACCAAAATGTTAATTATTTATTTAACCTGGTG
GGAGGTGGTTTTCTTTCTTTCTTTCTTTTGGAGACAGAGTTTACGCTCTCGTGTGCCCAGGCCTGGAGTGG
AATGGCAAGCCGACAGANCCAGCACAGTGGTTANGATNAATTAAGCGGGCNAGTCCAATAAGTCTGAAGGCTGA
AACACACACACAGAGACGGTCACAAAGGGCGAAATNCGGCAAAATNCCATCCCCTGCGGGGCCGATCACAAAAT
GNATNGAGAAGGGGACCAAATGGCCAAAGANGAGCGAT

SEQ ID NO:224

>GL2_147_4_M13F Direction: sense

CTGCTTCTCTTCTAGGATAGTTTCCCTCTAGAAATCCATGTTGCTCCTTCTCTCATTAAATGATCAGATATTTAT
GTGCCTCCTTAGCTAAGAGGCCTTCCTTATCTACCCTAACTAAGAGAATCCCCATATCACTCTTATCCTTTTATC
CTGATTTATTTTTCCATAAACTTGCCACATTACTGGGCATTATCTATCATTTAGTAATGTGTATATTGCTTATT
GACAGTCTTCCCTACTATAAATTATGAGCTCCATGAGGCAGAGACTTGTGTTTGTTCACCTGCTGAATTCCCAGCAC
CTACAACAAGCCTACTCACTAAACATTTGTTGAATTAATCTTCAGGATACATCAAATGTAAAAATATCCTATCATA
TATTTATTGATGTGATGCATAGGAAAATGTCTGGAAGGGTACATACCAAAATGTTAATTATTTATTTAACCTGGT
GGGAGGTGGTAA

SEQ ID NO:225

>GL2_147_1_M13R Direction: sense

CTGCTTCTCTTAGGATAGTTTCCCTCTAGAAATCCATGTTGCTCCTTCTCTCATTAGATGATCAGATATTTAT
GTGCCTCCTTAGCTAAGAGGCCTTCCTTATCTACCCTAACTAAGAGAATCCCCATATCACTCTTATCCTTTTATC
5 CTGATTTATTTTCCATAAACTTGCCACATTACTGGGCATTATCTATCATTTAGTAATGTGTATATTGCTTATT
GACAGTCTTCCCTACTATAATATGAGCTCCATGAGGCAGAGACTTTGTTTTGTTCACTGCTGAATTCCCAGCACC
TACAACAAGCCTACTCACTAAACATTTGTTGAATTAATCTTCAGGATACATCAAATGTAAAAATATCCTATCATAT
ATTTATTGATGTGATGCATAGGAAAATGTCTGGAAGGGTACATACCAAATGTTAATTATTTATTTAACCTGGTG
10 GGAGGTGGTTTTTCTTTCTTTTCTTTNCTTATTGAGACAGAGTTTCGCTCTCGTTGCCCCGGCTGGAAGTGGA
TGGCAAGCCGGCAGGAACCCAAGACAACAAGGAGGTTAAGATCAGAATAAAAGACGGGGACCGCATCCAGAAAC

SEQ ID NO:226

>GL2_147_4_M13R Direction: anti-sense

CTTGCCATTCCACTCCAGCCTGGGCAACGAGAGCGAAACTCTGGTCTCAAAAAGAAAGGAAAAGGAAAAGAAAAC
15 CACCTCCCACCAGGGTTAAATAAATAAATTAACATTTTGGTATGTACCCTTCCAGGACATTTTCTATGCATCACA
TCAATAAATATATGATAGGATATTTACATTTGATGTATCCTGAAGATTAATTCAACAAATGTTTAGTGAGTAGG
CNTGTTAGTAGGTGCTGGGAATTCAGCAGTGAACAAAACAAAGTCTCTGCCTCATGGAGCTCATATTATAGTAGG
GAAGACTGTCAATAAGCAATATACACATTACTAAATGATAGATAATGCCAGTAATGTGGCAAGTTTTATGGAAA
AATNAATCAGGATAAAAGGATAAGAGTGATATGGGGATTCTNCTTAGATTAGGGTAGATAAAGGAANGCCTCTTA
20 GNCTNAAGAGGCCACATAAAACAAATCAATTCCTTGGGAAAATCCAANCTGNCTAAACGTAGACAGGGACAAGCA
ANAGGCGGACGG

SEQ ID NO:227

Clone 147 - Novel (Chromosome 17)>

CTGCTTCTCTTAGGATAGTTTCCCTCTAGAAATCCATGTTGCTCCTTCTCTCATTAAATGATCAGATATTTAT
25 GTGCCTCCTTAGCTAAGAGGCCTTCCTTATCTACCCTAACTAAGAGAATCCCCATATCACTCTTATCCTTTTATC
CTGATTTATTTTCCATAAACTTGCCACATTACTGGGCATTATCTATCATTTAGTAATGTGTATATTGCTTATT
GACAGTCTTCCCTACTATAATATGAGCTCCATGAGGCAGAGACTTTGTTTTGTTCACTGCTGAATTTCCAGCACC
TACAACAAGCCTACTCACTAAACATTTGTTGAATTAATCTTCAGGATACATCAAATGTAAAAATATCCTATCATAT
30 ATTTATTGATGTGATGCATAGGAAAATGTCTGGAAGGGTACATACCAAATGTTAATTATTTATTTAACCTGGTG
GGAGGTGGTTTTTCTTTCTTTTCTTTTCTTTTGGAGACAGAGTTTCGCTCTCGTTGCCCAGGCTGGAGTGGAAT
GGCAAG

SEQ ID NO:228

Novel (chromosome 8)

>GL1_41_2_G3F1 Direction: sense

GCTTGCAAAGGAGAGGCTGTGACTACCAAGGTCGTGTCAACAACCTGAATGGCTGAAATACCCAACTTGCCCATG
CAAATGGGCTTGGGTCTCTCCTGGCAGCCGCTTTGAAGGCTCTAGACTTATCTGTGAACCTCTTTTGGAGAGG
GTCTTTCCAAGTAGTGGTTTATTCTTTGACTCTCCTCATACCTTTTTTGGCAGAGAGTGAGAGTGAGAAGGGAGG
40 GCTAATGCCTGAGCTCCTGCCCTTTCTATGCAGTGAGGGTCAAGATCCTCAGCTAGTGTGTTGAGGGAAGTGGTGT
AACCTGGGTCTCTCATTCTTCTACCATCCAAGTTGCC

SEQ ID NO:229

>GL1_41_2_G3R1 Direction: anti-sense

GGCAACTTGGATGGTAGAAAATGAGAGACCCAGGTTACACCAGGTTCCCTCAAACACTAGCTGAGGATCTTGACC
CTCACTGCNTAGAAAGGGCAGGAGCTCAGGCATTAGCCCTCCCTTCTCACTCTCACTCTCTGGCAAAAAGGTAT
GAGGAGAGTCAAAGAATAAACCACTAGTTGGAAAGACCCTCTCAAAAAGGAGTTACAGATAAGTCTAGAGCCT
TCAAAGGCGGCTGCCAGGAGAGACCCAAGCCATTGCGATGGGCAAGTTTGGGTATTTAGCCATTAGTTGTTG
50 ACACGACTTGGTAGTCACAGCCTCTCCTTTGCAAGC

SEQ ID NO:230

>GL1_41_3_G3F1 Direction: sense

GCTTGCAAAGGAGAGGCTGTGACTACCAAGTCGTGTCAACAACCTGAATGGCTGAAATACCCAACTTGCCCATGC
AAATGGGCTTGGGTCTCTCCTGGCAGCCGCTTTGAAGGCTCTAGACTTATCTGTGAACCTCTTTTGGAGAGGG
55 TCTTTCCAAGTAGTGGTTTATTCTTTGACTCTCCTCATACCTTTTTTGGCAGAGAGTGAGAGTGAGAAGGGAGGG
CTAATGCCTGAGCTCCTGCCCTTTCTATGCAGTGAGGGTCAAGATCCTCAGCTAGTGTGTTGAGGGAAGTGGTGTA
GCCTGGGTCTCTCATTCTTCTACCATCCAAGTTGCC

SEQ ID NO:231

>GL1_41_3_G3R1 Direction: N/A

GGCAACTTGGATGGTAGAAAATGAGAGACCCAGGCTACACCAGNTTCCCTCAAACACTAGCTGAGGATCTTGACC
CTCACTGCATAGAAAGGGCAGGAGCTCAGGCATTAGCCCTCCCTTCTCACTCTCACTCTCTGGCAAAAAGGTAT
GAGGAGAGTCAAANAATAAACCACTANTTGGANAGACCCCTCTCAANAAAGGAGTTCACAGATAANTNTATAGCCT
TCAAAGGCGGCTGCCAGGAGAGAC

SEQ ID NO:232

Novel (chromosome 8)>

GCTTGCAAAGGAGAGGCTGTGACTACCAAGTCGTGTCAACAACCTGAATGGCTGAAATACCCAAACTTGCCCATGC
AAATGGGCTTGGGTCTCTCCTGGCAGCCGCCCTTTGAAGGCTCTAGACTTATCTGTGAACTCCTTTTTTGAGAGGG
TCTTTCCAACCTAGTGGTTTATTCTTTGACTCTCCTCATACCTTTTTTTGCCAGAGAGTGAGAGTGAGAAGGGAGGG
CTAATGCCTGAGCTCCTGCCCTTTCTATGCAGTGAGGGTCAAGATCCTCAGCTAGTGTGTTGAGGGAACCTGGTGTA
ACCTGGGTCTCTCATTTTCTACCATCCAAGTTGCC

SEQ ID NO:233

Novel (chromosome 9)

>GL2-49-5M13R Direction: anti-sense

CTGGCCGCACGCAGCTCTTTTCTGGGGATCCCAGATGGTTGTGCGCTGGGGGCCTGGCGGAGGTCCGAGTCGGG
TTGGATGGCTGTGTCCGTCGCCATGTTTGTGTGGTGCTGGCCCCGTGGTCCCATGTTCTGGCAGCTTCTGGCTC
CAGCCTCAGCCTGTAGTGGGGTGTACGTGGTAGNAGTGCCTGAGCTGGGGGCCTGCAGTGGGGTAGGTCACTG
GTAGGAGTGCAGNTAGCATAGAAGTTANAAAGGTTACTAGTAGAGACTGGTCCCATACTTAGAGATGGGACACCA
TGACCTGATGGNTACNACCCTGACAGGCCGAGAAAACGACACACTGGAGCCACCGACTGACACAACGGGCCNT
GGGNACGGACGTCCAATNGCGTATGANTACCANGACAGTCAGGANCCCCAATGACCCGGAAGCAGGAAACGGGG
CAGAAAAANCAATAGCCAAGGAATAGTCACANGGTGCACANAACCTGNGCAGGGACCAGNAAACCAGAANGACC
AATGCACANTACATAAGAAGGGGGGGGGCACNCCACAAAAGATATGCCGGGCGACGCCAGATAAGATAAGGGACT
AGGGAAGGAGTCCGGGNAAANATGTAACGAAAGAGTATCGAACC AATGGAGGGNCGCCGGGTACCCGNGCGNAAG
GACAAACAANAAGCAGGCATCCGGACTGGGAACCAATGGGGGGGAAAGAAGCAGCGCCCCCATGGGGANAGGGAT
GTAAANCCACAACAACACGTGTNAAACG

SEQ ID NO:234

>GL2-49-4M13F Direction: N/A

AATTGAANTTACACTCACTATAGGGCGAATTGGGCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGAAT
ATCTGCAGAATTGCCCTTCTCTCGGCATGGACGAGCTGTACAAGGAGGAGGCCGCAAGGCCGGTGGCAGCGGTG
GCTCCAAGTGGTGGCCTTGGGGGTTCTTGCGGCTGCAGGGAGACCACAGGCAGGTGCCACCTAAGAGGGACAGC
CACAGAAACCCTCTAACCTCAGCACTGCACTCCACCACGAACACCCACGCAGGCCCTCAGGCAACCCGGCAACC
TTCCAAACCAACGGAACACCCACACAGGCTGAGGCTGGAGCCAGAAGCCGCCAGAACATGGGACCACAGGGCCA
GGCACCACACAAACATGGCGACGGACACAGCCAATCCAAACCCCGGAAAAACCGATTTCCNCGNGGGGGGAAAA
ACCGCCCCAGTTTTCCACCCCGGGAGAGCGCCCCAAAAGGGGGGGGNACACCCACCCACNACCCCAAAGGGGN
CGCAGGGGCCAAACAGCAGAAANCCCCGAANTTCCCTTGCGGGGGGAATTGCGCCGCGCCAAAGGGGAACNAGCA
AGGAAGGCCATTGCCGGTTGCCGGGCCCAAGCACCGCCAGGAACAGCGCGCAAGGCGAC

SEQ ID NO:235

>GL2-49-1M13F Direction: sense

CTGCAGGGAGACCACAGGCAGGTGCCCACCTAAGAGGGACAGCCACAGAAACCCTCTAACCTCAGCACTGCACTC
CACCACGACCACCCACGCAGGCCCTCAGCACCGCACTCCACCACGAACACCCACACAGGCTGAGGCTGGAGCC
AGAAGCTGCCAGAACATGGGACCACAGGGCCAGGCACCACACAAACATGGCGACGGACACAGCCATCCAACCCGA
CTCGGACCTCCGCCAGGCCCCCAGCGCACAAACCATCTGGGATCCCCAGGAAAAGAGCTGCGTGCGGCCAG

SEQ ID NO:236

>GL2-49-1M13R Direction: anti-sense

GCTGGCCGCACGCAGCTCTTTTCTGGGGATCCCAGATGGTTGTGCGCTGGGGGCCTGGCGGAGGTCCGAGTCGG
GTTGGATGGCTGTGTCCGTCGCCATGTTTGTGTGGTGCTGGCCCTGTGGTCCCATGTTCTGGCAGCTTCTGGCT
CCAGCCTCAGCCTGTGTGGGGTGTTCGTGCTGGAGTGCGGTGCTGAGGGCCTGCGTGGGGTGGTCTGTGGTGGAGT
GCAGTGCTGAGGTTAGAGGGTTTCTGTGGCTGTCCCTCTTAGGTGGGCACCTGCCTGTGGTCTCCCTGCAG

SEQ ID NO:237

>GL2-49-2M13F Direction: sense

CTGCAGGGAGACCACAGGCAGGTGCCCACCTAAGAGGGACAGCCACAGAAACCCTCTAACCTCAGCACTGCACTC
CACCACGACCACCCACGCAGGCCCTCAGCACCGCACTCCACCACGAACACCCACACAGGCTGAGGCTGGAGCC

AGAAGCTGCCAGAACATGGGACCACAGGGCCAGGCACCACACAAACATGGCGACGGACACAGCCATCCAACCCGA
CTCGGACCTCCGCCAGGCCCCCAGCGCACAAACCATCTGGGATCCCCAGGAAAAGAGCTGCGTGCGGCCAG

SEQ ID NO:238

>GL2-49-2M13R Direction: N/A

CTGGCCGCACGCAGCTCTTTTCCTGGGGATCCCAGATGGTTGTGCGCTGGGGGCTGGCGGAGGTCCGAGTCGGG
TTGGATGGCTGTGTCCGTCGCCATGTTTGTGTGGTGCCCTGGCCCTGTGGTCCCATGTTCTGGCAGCTTCTGGCTC
CAGCCTCAGCCTGTGTGGGGTGTTCGTGGTGGAGTGCGGTGCTGAGGGCCTGCGTGGGGTGGTCTGTGGTGGAGTG
CAGTGCTGAGGTTANAGGGTTTCTGTGGCTGTCCCTCTTAGGTGGGCACCTGCCTGTGGTCTCCCTGCAGCCGCA
AAACCCANCACTGGAGCCACCGCTG

SEQ ID NO:239

>GL2-49-4M13R Direction: N/A

CTGGCCGCACGCAGCTCTTTTCCTGGGGATCCCAGATGGTTGTGCGCTGGGGGCTGGCGGAGGTCCGAGTCGGG
TTGGATGGCTGTGTCCGTCGCCATGTTTGTGTGGTGCCCTGGCCCTGTGGTCCCATGTTCTGGCGGCTTCTGGCTC
CAGCCTCAGCCTGTGTGGGGTGTTCGTGGTGGAGTGCGGTGCTGAGGGCCTGCGTGGGGTGGTCTGTGGTGGAGTG
CAGTGCTGAGGTTAGAGGGTTTCTGTGGCTGTCCCTCTTAGGTGGGCACCTGCCTGTGGTCTCCCTGCAG

SEQ ID NO:240

>GL2_72_G07_G3F1 Direction: N/A

CTGCAGGGAGACCCCGCAGGCAGGTGCCACCTAAGGGGGACAGCTACAGAAACCTCTAACTCAGCACTGNACTC
CACCACGACACCCACGCAGGCCCTCAGCACCGCACTCCACCACGAACAACCCACACAGGCTGAGTGCTGGAGCC
AGAAGCTGGCAGAACATGGGACCACAGGGGCCAGGCACCACACAAACATGGCGACGGACACAGGCATNCAAACCG
ACTCGGACCTCCGCCAGGNCCTCAGCGCACAAACATCTGGGATCCCCAGGAAAAGAGCTGCGTGCGGGCAGCGCA
GAAACCAGCACAGTGTTAGATATGATTAAGCGGGCGNGTCGANTAATCTGAGGTCTGATACTCACTGACTGTCTG
TAAGGGNGAATTTCGCGGGCGCGTAAATTCAATTCGGCCTATAGTGAGTCGTATTACAATTCAGTGGGCGGCGTTT
TACAACGTCTGTGACTGGGAAAAACCTGNGCGTTATCCAATTAATCTGNGCTTGGAGAATTTCCCTTTTGCAGA
CTGGGCGTAATAACGAAAAAGGNNCCGAA

SEQ ID NO:241

>GL2_72_G08_G3F1 Direction: sense

CTGCAGGGAGACACAGGCAGGTGCCACCTAAGAGGGACAGCCACAGAAACCTCTAACCTCAGCACTGCACTCC
ACCACGACCACCCACGCAGGCCCTCAGCACCGCACTCCACCACGAACACCCACACAGGCTGAGGCTGGAGCCA
GAAGCTGCCAGAACATGGGACCACAGGGCCAGGCACCACACAAACATGGCGACGGACACAGCCATCCAACCCGAC
TCGGACCTCCGCCAGGCCCCCAGCGCACAAACATCTGGGATCCCCAGGAAAAGAGCTGCGTGCGGCCAG

SEQ ID NO:242

>GL2_72_G09_G3F1 Direction: sense

CTGCAGGAGACCGCAGGCAGGTGCCACCTAAGGGGGACAGCTACAGAAACCTCTAACCTCAGCACTGCACTCC
ACCACGAACACCCACGCAGGCCCTCAGCACCGCACTCCACCACGAACACCCACACAGGCTGAGGCTGGAGCCA
GAAGCTGGCAGAACATGGGACCACAGGGGCCAGGCACCACACAAACATGGCGACGGACACAGCCATCCAACCCGA
CTCGGACCTCCGCCAGGCCCCCAGCGCACAAACATCTGGGATCCCCAGGAAAAGAGCTGCGTGCGGCCAG

SEQ ID NO:243

>GL2_95_2B10_G3F1 Direction: sense

CTGCAGGAGACACAGGCAGGTGCCACCTAAGAGGGACAGCCACAGAAACCTCTAACCTCAGCACTGCACTCCA
CCACGACCACCCACGCAGGCCCTCAGCACCGCACTCCACCACGAACACCCACACAGGCTGAGGCTGGAGCCAG
AAGCTGCCAGAACATGGGACCACAGGGCCAGGCACCACACAAACATGGCGACGGACACAGCCATCCAACCCGACT
CGGACCTCCGCCAGGCCCCCAGCGCACAAACATCTGGGATCCCCAGGAAAAGAGCTGCGTGCGGCCAG

SEQ ID NO:244

>GL2_95_2B11_G3F1 Direction: sense

CCGCAGTGGAGACACAGGCAGGTGCCACCTAAGAGGGACAGCCACAGAAACCTCTAACCTCAGCACTGCACTC
CACCACGACACCCACGCAGGCCCTCAGCACCGCACTCCACCACGAAAAACCCACACAGGCGTGAGGCTGGAGCCAG
AAGCTGCCAGAACATGGGANACAGGGNCAGGCGACCACACAAACATGGCGACGTGACACAGGCATCCAAACCGA
CATCGGACTCCGCAGGGCCCCCAGCGCACAAACATACTGGGATCCCCACGAAAAGAGCTGCGTGCTGGGCAGNAG
CGCNAGAAACCAGCACAGCTGGTTAGAGTACAGATAAANACGGGCGCGTCGCACGTAAANCTGAAAAAGGTCGTG
AGTATAATCAACATGTANACTGNTT

SEQ ID NO:245

>GL2-49-5M13F Direction: sense

CTGCAGGGGAGACCACAGGCAGGTGCCCACCTAAGAGGGACAGCCACAGAAACCTCTAACCTCAGCACTGCACT
CCACCACGANACCCCCACGCAGGCCCCCAGCACCGCACTCCACCACGAAACACCCCAACACAGGCTGAGGCTGGA
GCCAGAAGCTGGCCAGAACATGGGACCAACGNGCGGNNCCCCAACGCNNGGCGGGCCAGAACGACCGAAAAGCG
AAACCCCAAAAGAAGNNGCGCAAATTATGCAGACCAGGAAACAGATGAAAACAAACCAAGGCGACCAATACC
CCAAAANANCACCGCNCAGGAAAGCATACACGCGGGGAAACGCCCAGTTCCACCACAGGGGCCCCACAAGGAGGG
GACCCNCCACCNACGNGCAGGGGCACCGGCCAGCAGACAGCGAGAGCGCCGCGAAGTACCGATTAGGAGGCGGCG
AATACCCGACCACCAAAAAAGGGGGAAACACGAAAGAAAGCTGGGAAAACGGCCCGATTTTAGGACCACGGGAGT
TTGACCGCACGGGGAGACGCCCCCAGAGGGACACCCCCAGAGGCACAAAAAGCAAGAAGACACCCCCCCCCAA
AGAGACCGCTCAAGACCCGAAAGAGGTTGGAGCGGCTTCTTAACAGAGGAAGCTTAACNCGGCACAGTTNAAAC
GAAGGCGAGTGGGC

SEQ ID NO:246

>GL2_49_1_M13R Direction: anti-sense

CTGGCCGCACGCAGCTCTTTTCTGGGGATCCAGATGGTTGTGCGCTGGGGGCTGGTGGAGGTCCGAGTCGGG
TTGGATGGCTGATGTCCGCCGCCATGATTTGATAGATGGATGCCTGGACCCTGATGAGTCCCAATGATTCCTGGCA
GACATTCTGGACATCCAGCACATCAGACCTGATGATGGGGATGNTTCGATGGATG

SEQ ID NO:247

>GL2_49_2_M13R Direction: N/A

CTGGCCGCACGCAGCTCTTTTCTGGGGATCCAGATGGTTGTGCGCTGGGGGCTGGCGGAGGTCCGAGTCGGG
TTGGATGGCTGGTTCGCGCAGTGGTGGTGCCTGGCCTGGGCTGTCTGGCAGCTTCTGGCTCCAGCCTCAGCCCTG
TGTGGGGCTGGTTCCGNTGGTGGAGTGCGGGTGCTGAAGGCCATGCGATGGAGAGTGGATCNAGGAAGGAGATGC
AGAAGCATAGAGAGTTACGAGGGTATCTAGTGGCTGTACCTCTATAGGTAGGGCACCATGACACTAGATGGTAC
TNCCTTGACAGAACGCAAGAAACACCAAGAAAAACCTGAGAGGCCAACCAGAAATGAC

SEQ ID NO:248

Clone 49 -Novel (Chromosome 9)>

CTGCAGGAGACCACAGGCAGGTGCCCACCTAAGAGGGACAGCCACAGAAACCTCTAACCTCAGCACTGCACTCC
ACCACGACCACCCACGCAGGCCCTCAGCAGCGCACTCCACCACGAACACCCACACAGGCTGAGGCTGGAGCCA
GAAGCTGCCAGAACATGGGACCACAGGGCCAGGCACCACACAAACATGGCGACGACACAGCCATCCAACCCGAC
TCGGACCTCCGCCAGGCCCCAGCGCACAAACCTCTGGGATCCCCAGGAAAAGAGCTGCGTGCGGCCAG

SEQ ID NO:249

Novel (clone 1/226)

>GL2_1_3_M13F Direction: anti-sense

ACTAGTTCTAGATCGCGAGCGGCGCCCTTTTTTTTTTTTTTTTAAAAAGTCTATATTTTTATATTGGGGGGAGG
GAGTAGAAAAGCAAGCCCCCTATACTGGGCCCTATTAGTGGCAGCTTCTTGTTCATAGGATTAAGGAAGACTCT
GAGGAAATAAAAGTTGTTTGGAAAAATCCAGGTGTAGTTGCTTTGTATGTTGTGATGGGTAGAAGGGATGAAGTG
AAGTGTGAAGGCCCTCATAACCTCCATCTTGCTCAGGACTATAGTCTTGAAACCTNNGGGCGGAGAAAAGCGC
CAACATTTTCATNCCCTGCATACATAAGGGGAGAAGGAGACAGGACAACGATAAGTGAAAGAGAACAGAACGAAG
AAAAGAAGCGGANAAACGGCCCCAGACAATAGTAAGGGCGAANGAATGGGCAGAAAGNCTTGCAANCGTNCCCGGG
GCAATACAATCCCTTGAACACAAGANCCAGGGCANNCCGGGACCACGCAAGAAAACCAAGAAACACAT

SEQ ID NO:250

>GL2_81B_E02_G3F1 Direction: N/A

GCCCGAGCCGGACTGGTCAGGATGATCACCGGACGCGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCT
CTTCTTGCTTGTGCTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGTGAAATGAAAGTGGCGCTTTC
TCCGGGCCCAGGGTCCAGGACATAGTCTGAGTGCAAGATGGAGGGTATGAGGGGGCTTCACACTTCACTTCATC
CTTTTAACCATTAATAACAAAGCGAACTACANCTGGATTTTTCCAAACAAATTTTATTTCTCAGAGTCTTCCCT
TAATCCTATGGAACAAGAAGCTGGCACTGAATAGTGGGCCAGTATAGGGGCTTGCCTTTCTANATCCCTTCCCC
CAATATTAATAATATATGACTTTTAAAAAAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGG
CCAGACAGTTGGTTAGATAGATAAAGCGGCGCGTCTGACTANTCTGAGGTCTGATACTCACTGACTGTGCTAAGGG
CGAATTCTGTTTTAAACCTGCAGGACTAGTCCCTTTATGAGGGTTAATTCTGAGCTTGGCGTAATCAT

SEQ ID NO:251

>GL2-96-1M13R Direction: sense

GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCT
TCTTGCTTGTGCTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTCTCTC

GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACGTTCACTTCAGTCCCT
TCTACCCATCACAACATACAAGGCAACTACACCTGGATTTTTCCAAACAACTTTTATTTCTCAGAGTCTTCCTT
AATCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAAT
ATAAAATATAGACTTTTAAAAAAGGGCGCCGCTCGCGATCTAGAACTAGT

5 SEQ ID NO:252

>GL2-96-3M13F Direction: sense

GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCT
TCTTGCTTGTGCTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
10 GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACTTCACTTCATCCCTTC
TACCCATCACAACATACAAGCAACTACACCTGGATTTTTCCAAACAACTTTTATTTCTCAGAGTCTTCCTTAA
TCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATAT
AAAAATATAGAC

15 SEQ ID NO:253

>GL2-96-3M13R Direction: N/A

ACTAGTTCTAGATCGCGAGCGGCGCCCTTTTTTTTTTTTTTAAAAAGTCTATATTTTTTATATTGGGGGGAGGG
AGTAGAAAAGCAAGCCCTATACTGGGCCCTATTCACTGGCAGCTTCTTGTTCCATAGGATTAAGGAAGACTCTG
AGGAAATAAAAGTTGTTTGAAAAATCCAGGTGTAGTTGCTTNTGNTATGTTGTGATG

20 SEQ ID NO:254

>GL2-96-4M13F Direction: anti-sense

ACTAGTTCTAGATCGCGAGCGGCGCCCTTTTTTTTTTTTTTAAAAAGTCTATATTTTTTATATTGGGGGGAGGGAGTAG
AAAAGCAAGCCCTATACTGGGCCCTATTCACTGGCAGCTTCTTGTTCCATAGGATTAAGGAAGACTCTGAGGAA
25 ATAAAAGTTGTTTGAAAAATCCAGGTGTAGTTGCTTTGTATGTTGTGATGGGTAGAAGGGATGAAGTGAAGTGT
GAAGGCCCTCATACCTCCATCTTGCTTCAGACTATGTCTGGAACCTGGGGCGGAGAAAGCGCCACTTTTCAT
TCCTGCTTCTTGGGATTGTTGACGGCCACGTAGTGATAGAGAACGACAAGCAAGAAGAGCGACACGCCCAGCATG
TTGGCGAAGATGGCGAGCTGCACGTCCGTGATCATCTGACCAGTCCGGCTCGGCCCGCAGAAACCAGCACACTG
GAGCCACCGCTGCCACCGGCCTT

30 SEQ ID NO:255

>GL2-96-5M13F Direction: sense

GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCT
TCTTGCTTGTGCTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
35 GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACTTCACTTCATCCCTTC
TACCCATCACAACATACAAGCAACTACACCTGGATTTTTCCAAACAACTTTTATTTCTCAGAGTCTTCCTTAA
TCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATAT
AAAAATATAGACTTTTAAAAAAGGGCGCCGCTCGCGATCTAGAACTAGTCCGCAAAACCCAGCAC
AGTGGTTAGATAGATAAAGC

40 SEQ ID NO:256

>GL2-96-5M13R Direction: N/A

ACTAGTTCTAGATCGCGAGCGGCGCCCTTTTTTTTTTTTTTAAAAAGTCTATATTTTTTATATTGGGGGGAGGGA
GTAGAAAAGCAAGCCCTATACTGGGCCCTATTCACTGGCAGCTTCTTGTTCCATAGGATTAAGGAAGACTCTGA
45 GGAAATAAAAGTTGTTTGAAAAATCCAGGTGTAGTTGCATATAAGTATGTTGTGATA

50 SEQ ID NO:257

>GL2_1_1_M13R Direction: sense

GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCT
TCTTGCTTGTGCTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
50 GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACTTCACTTCATCCCTTC
TACCCATCACAACATACAAGCAACTACACCTGGATTTTTCCAAACAACTTTTATTTCTCAGAGTCTTCCTTAA
TCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATAT
AAAAATATAGACTTTTAAAAAAGGGCGCCGATCGCGATCTAGAACTAGT

55 SEQ ID NO:258

>GL2_1_2_M13F Direction: sense

GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCT
TCTTGCTTGTGCTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
60 GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAAGGGCCTTCACACTTCACTTCATCCCTTC

TACCCATCACAACATACAAAGCAACTACACCTGGATTTTTCCAAACAACTTTATTTTCCTCAGAGTCTTCCTTAA
TCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATAT
AAAATACTAGACTTATT

5 SEQ ID NO:259

>GL2_1_2_M13R Direction: anti-sense

ACTAGTTCTAGATCGCGAGCGGCGCCCTTTTTTTTTTTTTTTTTTAAAAAGTCTATATTTTATATTGGGGGGAGG
GAGTAGAAAAGCAAGCCCTATACTGGGCCCTATTCAGTGGCAGCTTCTTGTTCCATAGGATTAAGGAAGACTCT
10 GAGGAAATAAAAGTTGTTTGGAAAAATCCAGGTGTAGTTGCTTTGTATGTTGTGATGGGTAGAAGGGATGAAGTG
AAGTGTGAAGGCCCTTCATACCCTCCATCTTGCCTCAGACTATGTACCTGGAACCCTGGGGCNGAGAAAGCGCCA
CNTTCATTCTGCTTCTAGGGATCGNNAGACGGACACGATAGNTGACTAGAGAACGACAAGCAAGAA

SEQ ID NO:260

>GL2_226_1_1_M13F Direction: sense

15 GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCT
TCTTGCTTGTGCTTCTCTATCACCACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCCTTCACACTTCACTTCATCCCTTC
TACCCATCACAACATACAAAGCAACTACACCTGGATTTTTCCAAACAACTTTATTTTCCTCAGAGTCTTCCTAAT
20 CCTCATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATAT
AAAATATAGACTTTTAAAAAAGGGCGCCGCTCGCGATCTAGAACTAGT

SEQ ID NO:261

>GL2_226_1_1_M13R Direction: anti-sense

25 ACTAGTTCTAGATCGCGAGCGGCGCCCTTTTTTTTTTTTTTTTTTAAAAAGTCTATATTTTATATTGGGGGGAGGGA
GTAGAAAAGCAAGCCCTATACTGGGCCCTATTCAGTGGCAGCTTCTTGTTCCATAGGATTAGGAAGACTCTGAG
GAAATAAAAGTTGTTTGGAAAAATCCAGGTGTAGTTGCTTTGTATGTTGTGATGGGTAGAAGGGATGAAGTGAAG
TGTCAGGGCCCTCATACCCTCCATCTTGCCTCAGACTATGTCCTGGAACCCTGGGGCGGAGAAAGCGCCNCTTT
CATTCCTGCTTCTTGGGATTGTTGACGGCCACGTGGTGATAGAGAACGACAAGCAAGAAGAGCGANACGCCACG
30 ATGTTGGCGAAGATGGCGAGCTGCACGTCCGTGATCATCCTGACCAGTCCAGGCTCGGC

SEQ ID NO:262

>GL2_226_1_2_M13F Direction: sense

35 GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCT
TCTTGCTTGTGCTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCCTTCACACTTCACTTCATCCCTTC
TACCCATCACAACATACAAAGCAACTACACCTGGATTTTTCCAAACAACTTTATTTTCCTCAGAGTCTTCCTTAA
TCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATAT
AAAATATAGACTTTTAAAAAAGGGCGCCGCTCGCGATCTAGAACTAGT

40 SEQ ID NO:263

>GL2_226_1_2_M13R Direction: anti-sense

45 ACTAGTTCTAGATCGCGAGCGGCGCCCTTTTTTTTTTTTTTTTTTAAAAAGTCTATATTTTATATTGGGGGGAGGGA
GTAGAAAAGCAAGCCCTATACTGGGCCCTATTCAGTGGCAGCTTCTTGTTCCATAGGATTAAGGAAGACTCTGA
GGAAATAAAAGTTGTTTGGAAAAATCCAGGTGTAGTTGCTTTGTATGTTGTGATGGGTAGAAGGGATGAAGTGAA
GTGTGAAGGCCCTCATACCCTCCATCTTGCCTCAGACTATGTCCTGGAACCCTGGGGCGGAGAAAGCGCCACTT
TCATTCTGCTTCTTGGGATTGTTGACGGCCACGTAGTGATAGAGAACGACCAGCAAGAAGAGCGACACGCCA
GCATGTTGGCGAAGATGGCGAGCTGCACGTCCGTGATCATCCTGACCAGGTCCGGCATCGG

SEQ ID NO:264

>GL2_226_1_3_M13F Direction: sense

50 GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCT
TCTTGCTTGTGCTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCCTTCACACTTCACTTCATCCCTTC
TACCCATCACAACATACAAAGCAACTACACCTGGATTTTTCCAAACAACTTTATTTTCCTCAGAGTCTTCCTTAA
55 TCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATAT
AAAATATAGACTTTTAAAAAAGGGCGCCGCTCGCGATCTAGAACTAGT

SEQ ID NO:265

>GL2_226_1_4_M13F Direction: sense

5 GCGGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCGTCGCCAACATGCTGGGCGTGTGCTCTCT
TCTTGCTTGTCGTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACTTCACCTTCATCCCTTC
TACCCATCACAACATACAAAGCAACTACACCTGGATTTTTCCAAACAACCTTTATTTCTCAGAGTCTTCCTTA
ATCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATA
TAAAAATATAGACTTTTAAAAAAGGGCGCGCTCGCGATCTAGAACTAGT

10 SEQ ID NO:266

>GL2_81B_E01_G3F1 Direction: sense

15 GCGGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCTT
CTTGCTTGTCGTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAGGTGGCGCTTTCTCCG
CCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACTTCACCTTCATCCCTTCT
ATCCATCACAACATACAAAGCAACTACACCTGGATTTTTCCAAACAACCTTTATTTCTCAGAGGTCTTCCTTA
ATCCTATGGAACAAGAAGCTGNCACCTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATA
TNAATATAGACTTTTAAAAAAGGGCGCGCTCGCGATCTAGAACTAGT

20 SEQ ID NO:267

>GL2_81B_E05_G3F1 Direction: N/A

25 GCGGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCTT
TCTTGCTTGTCGTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
GNCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACTTCACCTTCATCCCTTC
TACCCATCACAACATACAAAGCAACTACACCTGGATTTTTCCAAACAACCTTTATTTCTCAGAGTCTTCCTTAA
TCCTATGGAACAAGAAGCTGGCACTGAATAGTGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATA
TNAATATAGACTTTTAAAAAAGGGGNGCGGTGCGATCTAGAACTAGTCCGNNAGAAACCAGCA
CAGTGGTTAGATAGATAAAGCGGGCGGTGCGACTANTCTGAGGTCTGATACTACTGACT

30 SEQ ID NO:268

>GL2_81B_E06_G3F1 Direction: sense

35 GCGGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCTT
TCTTGCTTGTCGTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACTTCACCTTCATCCCTTC
TACCCATCACAACATACAAAGCAACTACACCTGGATTTTTCCAAACAACCTTTATTTCTCAGAGTCTTCCTTAA
TCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATAT
AAAAATATAGACTTTTAAAAAAGGGGCGCGCTCGCGATCTAGAACTAGTCCGNAGAACCCAG
CACAGTGGTTAGATAGATAAAGCGGCCGCTCGACTAGTCTGAGGTCTGATACTACTGACTGTGTAAGGGCGAA
TTCGTTTAAACCTGCAGGACTAGTCCCTTTATGAGGGTTAAATTCTGAGCTTGGCGTAATCATGGTCAT

40 SEQ ID NO:269

>GL2_1_3_M13R Direction: sense

45 GCGGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCTT
TCTTGCTTGTCGTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACTTCACCTTCATCCCTTC
TACCCATCACAACATACAAAGCAACTACACCTGGATTTTTCCAAACAACCTTTATTTCTCAGAGTCTTCCTTA
ATCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCAATTTTCTACTCCCTCCCCCAAA
TATAAAATATAGACTTTTAAACAAACACAACAACAACAAAAAGGGCGCCAGCTCGGCGATCTAGAACTAGT
CCGGCAAGAACCCAGNAACAGGGGTTAGATA

50 SEQ ID NO:270

>GL2_81B_E04_G3F1 Direction: N/A

55 GCGGAGCCGGACTGGTCAGGATGATCACGGACGCGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCTT
TCTTGCTTGTCGTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAAGTGGCGCTTTCTCC
GCCCCAGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCTTCACACTTCACCTTCATCCCTT
CTANCCATCACAACATACAAAGCAACTACACCTGGATTTTTCCAAACAACCTTTATTTCTCAGAGTCTTCCTTA
ATCCTATGGAACAAGAAGCTGGCACTGAATAGGGCCAGTATANGGGCTTGCGTTTTTCTACTCCCTCCCCCAAT
ATAAAATATAGACTTTTAAAAAATNAAGGGNGCGCTCGCGATACTAGAACTAGT

SEQ ID NO:271

>GL2_226_1_3_M13R Direction: anti-sense

CTAGTTCTAGATCGCGAGCGGCGCCCTTTTTTTTTTTTTTTTAAAAAGTCTATATTTTTATATTGGGGGGAGGGAG
TAGAAAAGCAAGCCCCCTATACTGGGCCCCTATTCACTGGCAGCTTCTTGTTCATAGGATTAAGGAAGACTCTGAG
GAAATAAAAGTTGTTTGGAAAAATCCAGGTGTAGTTGCTTTGTATGTTGTGATGGGTAGAAGGGATGAAGTGAAG
TGTGAAGGCCCCCTCATACCCTCCATCTTGCCTCAGACTATGTCTTGGAACCCCTGGGGCGGAGAAAGCGCCACTGT
TCATTCCTGCTNTCTTGGGATTGTTTCGACGGNCACGTAGATGATAGAGAACGACAAGCAAGAAGAGCGAANACGC
CCAGCATGTAAGGCGAAGATGGCGAGCTGCACGTCCGTGATCATCTTGACCAGTCCGGCTCGGA

SEQ ID NO:272

>GL2_226_1_4_M13R Direction: anti-sense

ACTAGTTCTAGATCGCGAGCGGCGCCCTTTTTTTTTTTTTTTTAAAAAGTCTATATTTTTATATTGGGGGGAGGGGA
GTAGAAAAGCAAGCCCCCTATACTGGGCCCCTATTCACTGGCAGCTTCTTGTTCATAGGATTAAGGAAGACTCTGA
GGAAATAAAAGTTGTTTGGAAAAATCCAGGTGTAGTTGCTTTGTATGTTGTGATGGGTAGAAGGGATGAAGTGAAG
GTGTGAAGGCCCCCTCATACCCTCCATCTTGCCTCAGACTATGTCTTGGAACCCCTGGGGCGGAGAAAGCGCCACTT
TCATTCCTGCTTCTTGGGATTGTTGACGGCCACGTAGTGATAGCGAACGACAAGCAAGAAGAGCGACACGCCAG
CATGTTGGCGACGATGGCGAGCTGCACGTCCGTGATCATCTTGACCAGTCCGGCTCGGCCCCGAAGAAGAGGAGGAG
ACACTGGAGCCACCGTGCACCGGCCCTTGGCGGCTACCTCCTGTACAGCTCGGTCCATGCCGAGAGAAGGGCG
AAATACATGCAGATATCCATCACAACCTGGCGGGCCGATCCGAGCATGCATCTAGAAGGGCCAAATACGCCATATA
GGNGAGTCGGAATAACAATTACCTGGCGGGCGATCCAAACGTCGTGACTGGAAAAAACCGGGGGATNACCCAA
AAAAGTTAAATCGGCCNGAAGGACANAACCCACGATTAGGCAANAGNGCGAAAAATAACNAAAA

SEQ ID NO:273

>GL2_81B_E03_G3F1 Direction: sense

GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCT
TCTTGCTTGTGCTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAGGTGGCGCTTTCTCC
GCCCCAGGGGTTCCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCCTTCACACTTCACTTCATCCCTT
CTACCCATCACAACATACAAAGCAACTACACCTGGATTTTCCAAACAACCTTTATTTCTCTCAGAGTCTTCCTTA
ATCCTATGGAACAAGAAGCTGCCACTGGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATA
TAAAAATATAGACTTTTAAAAAAGGGCGCCGCTCGCGATCTAGAAGTAGT

SEQ ID NO:274

Clone 1/226 - Novel (Clone1/226) >

GCCGAGCCGGACTGGTCAGGATGATCACGGACGTGCAGCTCGCCATCTTCGCCAACATGCTGGGCGTGTGCTCT
TCTTGCTTGTGCTTCTCTATCACTACGTGGCCGTCAACAATCCCAAGAAGCAGGAATGAAGGTGGCGCTTTCTCC
GCCCCAGGGTTCAGGACATAGTCTGAGGCAAGATGGAGGGTATGAGGGGCCCTTCACACTTCACTTCATCCCTTC
TACCCATCACAACATACAAAGCAACTACACCTGGATTTTCCAAACAACCTTTATTTCTCTCAGAGTCTTCCTTAA
TCCTATGGAACAAGAAGCTGCCACTGAATAGGGCCAGTATAGGGGCTTGCTTTTCTACTCCCTCCCCCAATAT
AAAAATATAGACTTTTAAAAAAGGGCGCCGCTCGCGATCTAGAAGTAGT

SEQ ID NO:275

Novel (FLJ10688)

>GL2_51_C08_G3F1 Direction: sense

CACCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCAGAAGCTGCCAGCACATCTCTGCCTCAGA
AGCCACACTTGAAGTTAGCACGCGTTAGAGTCAAAATGGCATAGTACTGTCTGAGGTGTCTGAGGTGGATC
GAAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACAAT
GGAAAAAGATTGGGGAAGTCAAGGCACCTCCCTTGCCCCATGGCATGTACTCTCACCAGTTTGTATCTGGTAGC
AAATACTACTTTG

SEQ ID NO:276

>GL2_51_C10_G3F1 Direction: sense

CACCCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCAGAAGCTGCCAGCACATCTCTGCCTCAG
AAGCCACACTTGAAGTTAGCACGCGTTAGAGTCAAAATGGCATAGTACTGTCTGAGGTGTCTGAGGTGGAT
CGAAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACA
TGGAAAAAGATTGGGGAAGTCAAGGCACCTCCCTTGCCCCATGGCATGTACTCTCACCAGTTTGTATCTGGTAGC
AAATACTACTTTG

SEQ ID NO:277

>GL2_51_C11_G3F1 Direction: N/A

CACCCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCCAGAAGCTGCCAGCACATCTCTGCCTCAGAA
AGCCACACTTGAAGTTAGCACGCGTTTCAGGGGTCAAAATGGCATAGTACTGTTCATGGAGTGTCTTGGAGGTGGATCG
CGAAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACAA
TGGAAAAAGATTGGGGAAGTCAAGGCACTTCCCTTGCCCATGGCATGTACTCTCACCAGTTTGTATCTGGTAGC
AAATACTACGTTTG

SEQ ID NO:278

>GL2_59_E07_G3F1 Direction: sense

CACCAGCACCTTACAGAAGCTCCACAACCACAGCGTCTGCCCCCAGAAGCTGCCAGCACATCTCTGCCTCAGAAA
CCACACTTGAAGTTAGCACGCGTTTCAGAGTCAAAATGGCATAGTACTGTTCATGGAGTGTCTTGGAGGTGGATCG
AAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACAA
GAAAAAGATTGGGGAAGTCAAGGCACTTCCCTTGCCCATGGCATGTACTCTCACCAGTTTGTATCTGGTAGCAA
ATACTACTTTG

SEQ ID NO:279

>GL2_59_E08_G3F1 Direction: sense

CACCCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCCAGAAGCTGCCAGCACATCTCTGCCTCAG
AAACCACACTTGAAGTTAGCACGCGTTTCAGAGTCAAAATGGCATAGTACTGTTCATGGAGTGTCTTGGAGGTGGATCG
CGAAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACAA
TGGAAAAAGATTGGGGAAGTCAAGGCACTTCCCTTGCCCATGGCATGTACTCTCACCAGTTTGTATCTGGTAGC
AAATACTACTTTG

SEQ ID NO:280

>GL2_59_E09_G3F1 Direction: sense

CACCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCCAGAAGCTGCCAGCACATCTCTGCCTCAGAA
GCCACACTTGAAGTTAGCACGCGTTTCAGAGTCAAAATGGCATAGTACTGTTCATGGAGTGTCTTGGAGGTGGATCG
AAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACAA
GAAAAAGATTGGGGAAGTCAAGGCACTTCCCTTGCCCATGGCATGTACTCTCACCAGTTTGTATCTGGTAGCAA
ATACTACTTTG

SEQ ID NO:281

>GL2_59_E10_G3F1 Direction: sense

CACCCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCCAGAAGCTGCCAGCACATCTCTGCCTCAGAA
AGCCACACTTGAAGTTAGCACGCGTTTCAGAGTCAAAATGGCATAGTACTGTTCATGGAGTGTCTTGGAGGTGGATC
GAAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACAA
GGAAAAAGATTGGGGAAGTCAAGGCACTTCCCTTGCCCATGGCATGTACTCTCACCAGTTTGTATCTGGTAGCA
ATACTACTTTG

SEQ ID NO:282

>GL2_59_E11_G3F1 Direction: sense

CACCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCCAGAAGCTGCCAGCACATCTCTGCCTCAGAA
AACCACACTTGAAGTTAGCACGCGTTTCAGAGTCAAAATGGCATAGTACTGTTCATGGAGTGTCTTGGAGGTGGATC
GAAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACAA
GGAAAAAGATTGGGGAAGTCAAGGCACTTCCCTTGCCCATGGCATGTACTCTCACCAGTTTGTATCTGGTAGCA
ATACTACTTTG

SEQ ID NO:283

>GL2_59_E12_G3F1 Direction: sense

CACCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCCAGAAGCTGCCAGCACATCTCTGCCTCAGAA
GCCACACTTGAAGTTAGCACGCGTTTCAGAGTCAAAATGGCATAGTACTGTTCATGGAGTGTCTTGGAGGTGGATCG
AAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACAA
GAAAAAGATTGGGGAAGTCAAGGCACTTCCCTTGCCCATGGCATGTACTCTCACCAGTTTGTATCTGGTAGCAA
ATACTACTTTG

SEQ ID NO:284

>GL2_51_C09_G3F1 Direction: N/A

CACCAGCACCTTACCAGAAGCTCCACAACCACAGGCGTCTGGCCCCCAAGAACAAACAGGGGCCCTTTGTGGTCTG
TNCCCCACAAAAGGGGGCCCCCTTATAAACACCCATAAAATTTTTTACCCCTTTTNTCCCCCTTGGGGTGTCCCC
CCCCCCTTTTTTCCCCAAAAAAGGGGNCCCCAAACCAAGGGGGCCCCCCCCCTTTAAAAACCCATAAAACCCC
CCTTTTTTTTTTTTTTNGGGGGCCCCCAACCCAAAC

SEQ ID NO:285

Novel (FLJ10688) >

CACCCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCCAGAAGCTGCCAGCACATCTCTGCCTCAGA
AGCCACACTTGAAGTTAGCAGCGTTTACAGAGTCAAATGGCATAGTACTGTCATGGAGTGTCTGGAGGTGGATC
GAAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTACCATGAGGAACCCAGTGCCACTGTGCCCTCACAAT
GGAAAAAGATTGGGAAGTCAAGGCACTTCCCTTGCCCATGGCATGTACTCTACCCAGTTTGTATCTGGTAGCA
AATACTACTTTG

SEQ ID NO:286

>gi|8922591|ref|NM_018179.1| Homo sapiens hypothetical protein FLJ10688
(FLJ10688), mRNA

ATCTAAACAAGAAAGTAGTGAGAGTTTGCCAAAAGAAGCCTTTCTGGTCTCTCTGATGAAGAGGATATT
TCGGGTGAAAAAGATGAGTCTGAAGTTATATCGCAAAATGAAACGTGCTCTCCAGCAGAAGTAGAAAGTA
ATGAAAAAGGACAACAGACCTGAGGAAGAAGAGCAAGTAATACATGAAGATGATGAAAGACCTTCTGAGAA
AAATGAATTTTCTAGACGAAAACGTTCTAAATCAGAAGACATGGACAATGTACAGTCTAAACGTCGTCGA
TATATGGAAGAAGAATATGAGGCAGAATTTCAAGTAAAGATTACAGCCAAAGGAGACATTAACCAGAAAC
TTCAAAAGGTTATACAGTGGTTGCTGGAAGAAAAATTGTGTGCGCTGCAGTGTGCTGTATTTGATAAGAC
TTTGGCAGAAATTGAAAACACGAGTGGAAAAGATTGAATGTAACAAGAGGCATAAAACAGTTCTCACTGAA
CTACAGGCCAAGATAGCCAGGTAAACCAAACGCTTTGAAGCAGCCAAAGAAGATCTTAAGAAAAGACATG
AACATCCACCCAACCCACAGTATCACCAGGAAAAACTGTAAATGATGTCAACAGCAATAATAACATGTC
TTACAGAAATGCAGGCACAGTGAGACAGATGCTGGAGTCCAAAAGAAATGTAAGCGAGAGTGCACCACCA
TCCTTTCAAACCTCCTGTGAATACAGTATCTTCAACCAATCTTGTCACTCCTCCAGCAGTTGTGAGTAGTC
AACCTAAATTGCAGACTCCAGTGAATTCGGGTTCCCTCACAGCAACGTCAGTTCTTCTGCACCCAATAC
AGCTACTGTAGTTGCTACTACTCAGGTGCCTAGTGGAAATCCCCAGCCTACAATCTCTTTACAGCCTTTG
CCAGTGATTTTGCATGTACCTGTTGCAGTATCCTCCCAGCCTCAGCTTCTACAGAGCCATCCAGGGACTT
TGGTGACTAATCAACCATCTGGCAATGTTGAATTCATTTCTGTGCAAAGCCACCTACAGTGAGTGGTCT
TACCAAAAATCCAGTATCCTTGCCATCCTTGCCAAATCCCACTAAACCAAAACAGTTCCCTTCTGTGCCC
AGTCCTAGTATTCAAAGGAACCTACTGCCAGTGTGTCACCATTGGGAACAACACTTGCTGTGCAGGCTG
TTCCAACAGCACACTCTATTGTACAAGCCACAAGGACTTCTTTACCCACAGTGGGGCCCATCAGGACTCTA
TAGTCCATCAACTAATCGAGGTCTTATACAGATGAAAATTCCAATTTCTGCAATTTAGTACTTTCGTCTGCT
GCAGAACAGAACAGCAATACCACCCCAAGAATTGAAAACAGACAAAACAAAACAAATAGATGCTTCTGTCA
GTAAGAAAGCAGCTGATAGCACATCACAGTGTGGAAGGCCACTGGCAGTGATTCAAGTGGTGTCAATTGA
TCTCACAATGGATGATGAAGAGAGTGGAGCTTCAACAAGACCCCAAAAACTAAATCACACTCCTGTATCA
ACCATGAGTTCTTCTCAGCCTGTGTACGACCATTGCAACCCATACAAACCAGCACCGCCTCTTCAACCAT
CTGGGGTGCCAACAAGTGGACCATCTCAGACCACCATACACTTACTACCTACAGCTCCAACCTACCGTGAA
TGTAACACATCGTCCAGTAACTCAGGTGACCACAAGACTCCCTGTACCAAGAGCTCCTGCAAACCACCAG
GTGGTTTATACAACCTCTTCTGCAACCACCAGCTCAGGCTCCCTTGCGAGGAACGTGTTATGCAGGCTCCTG
CTGTTTCGGCAGGTCAATCCCCAAATAGTGTTACAGTTCGAGTGCCTCAAACAACCACATATGTTGTAAA
CAATGGACTAACCCTGGGATCAACAGGACCTCAGCTCAGTGTCATCACCAGCACCAAGTGCATACT
GAGCCCCCAGCCCCGTGCACCCAGCACCTTACCAGAAGCTCCACAACCACAGCGTCTGCCCCCAGAAG
CTGCCAGCACATCTCTGCCTCAGAAGCCACACTTGAAGTTAGCACGCGTTCAGAGTCAAAATGGCATAGT
ACTGTCATGGGGTGTCTGGAGGTGGATCGAAGCTGTGCCACTGTTGATAGCTACCATCTCTATGCTTAC
CATGAGGAACCCAGTGCCACTGTGCCCTCACAATGGAAGAAAGATTGGGGAAGTCAAGGCACTTCCCTTGC
CCATGGCATGTACTCTACCCAGTTTGTATCTGGTAGCAATACTACTTTGTCAGTACGAGCCAAGGATAT
TTATGGACGTTTGGGCCTTTCTGTGATCCTCAGTCAACAGATGTGATCTCTTCTACCCAGAGCAGTTAA
ACCTTGAGGCCTTTATATTTTCTCTTTTAAATTTCCACCTTTTGGTCTTGTGTTTTAATCTTGTGCATG
ATACCCCATGTAAAATCCACCTTGTGCAAGATTTCTTGGACAGATGTGTGTATACACTACATTTGTTTTAT
AACCAGAAGCAAAATAAATCAGCCCACAAAGCT

SEQ ID NO:287

>gi|8922592|ref|NP_060649.1| hypothetical protein FLJ10688; likely ortholog of mouse activating transcription factor 7 interacting protein [Homo sapiens]

5 MDNVQSKRRRYMEEYEAEFQVKITAKGDINQKLQKVIQWLLKEKLCALQCAVFDKTLAELKTRVEKIEC
NKRHKTVLTELQAKIARLTKRFEAAKEDLKKRHEHPNPPVSPGKTVDVNSNNMSYRNAGTVRQML
KRNVSSEAPPSFQTPVNTVSSTNLVTPPAVVSSQPKLQTPVTSGSLTATSVLPAPNTATVVATTQVPSGN
10 PQPTISLQPLPVILHVPVAVSSQPQLLQSHPGTLVTNQPSGNVEFISVQSPPTVSGLTKNPVSLPLPNP
TKPNNVSPVSPSIQRNPTASAAPLGTTLAVQAVPTAHSIVQATRTSLPTVGPGLYSPSTNRGPIQMKI
PISAFSTSSAAEQNSNTTPRIENQTNKTIDASVSKKAADSTSQCCKATGSDSSGVIDLTMDDEESGASQD
PKKLNHTPVSTMSSSQPVSRPLQPIQPAPPLQPSGVPTSGPSQTTIHLLPTAPTTVNVTHRPVTQVTTTL
PVPRAPANHQVYVYTTLPAPPAQAPLRGTVMQAPAVRQVNPQNSVTVRVPQTTTYVNVNGLTLGSTGPQLT
VHHRPPQVHTEPPRPVHPAPLPEAPQQRLLPPEAASTSLPQKPHLKLARVQSQNGIVLSWGVLEVDRSCA
15 TVDSYHLAYAYHEEPSATVPSQWKKIGEVKALPLPMACLTLTQFVSGSKYYFAVRAKDIYGRFGFPFCDPQST
DVISSTQSS

SEQ ID NO:288

Novel (KIAA0217)

>G3_4_03_PCR_G3F1 Direction: N/A

20 TCTTGCGGCAGGGTCAAAACAACATCATGGTACAACCTTCTGGTAACATGTCTAGTGGGCAACAAAGTGTATGTTTT
CATGTAGTTCTATTACAAATTGGTTCTTTGAATTTATATTACATAATACCACTTGAACTGGAGTGAGAAGCTG
GTTACTTGCTTGCACTGCATCAGAACCAAGGTTTCTGGGGTGGCCCCCTGCAGTTCACTCCGCCAGGAGATCAAA
AAGGATGCGCTCTCTACATTCTTAAATGCCACAGTCCCCACCAACGGCTTAGTTGAAGCACAGCTCCCATCA
GCACCTATTCCCTGGCGCGGCCAGCACTGCCTCTACGAGGCCACACCTGCCTCTGCCAAGGCCACTGGCCCTCAT
25 TCATCTTCGGGGACGTTCTAGAGATATGCACTTACAG

SEQ ID NO:289

>GL3_4.03_A_M13F Direction: anti-sense

30 CTGTAAGTGCATATCTCTAGAACGTCCCCGAGATGAATGAGGGCCAGTGGCCTTGGCAGAGGCAGGTGTGGCCT
CGTAGAGGCAGTGCTGGCCGCGCCAGGGCATCAGTGCTGATGTGGGAGCTGTGCTTCCACCTAAGCCGTTGGTAG
GGGACTGTGGCATTAAAGAATGTAGAGAGCGCATCCTTTTGATCTCCTGGGCGGAGTGAACCTGCAGGGGCCAC
CCCAGAAACCTTGGTTCTGATGCACTGCAAGCAAGTAACCAGCTTCTCACTCCAGTTTCAAGTGGCTATTATGTA
ATATAAATTCAAAGCACATTGTGAATAGAACCTACATGAAAACATACACTTTGTTGCCCACTGACATGTTACCAG
35 AAGTTGTACCATGATGTTGTTTGTACCCCTG

SEQ ID NO:290

>GL3_4.03_A_M13R Direction: sense

40 CAGGGGTCAAACAACATCATGGTACAACCTTCTGGTAACATGTCTAGTGGGCAACAAAGTGTATGTTTTTCATGTAG
GTTCTATTACAAATGTGCTTTGAATTTATATTACATAATAGCCACTTGAACTGGAGTGAGAAGCTGGTTACTTG
CTTGCACTGCATCAGAACCAAGGTTTCTGGGGTGGCCCCCTGCAGGTTCACTCCGCCAGGAGATCAAAAAGGATG
CGCTCTCTACATTCTTAAATGCCACAGTCCCCACCAACGGCTTAGGTGGAAGCACAGCTCCCATCATCAGCACTG
ATGCCCTGGCGCGGCCAGCACTGCCTCTACGAGGCCACACCTGCCTCTGCCAAGGCCACTGGCCCTCATTCATCT
GCGGGGACGTTCTAGAGATATGCACTTACAG

SEQ ID NO:291

>GL3_4.03_B_M13F Direction: sense

45 CAGGGGTCAAACAACATCATGGTACAACCTTCTGGTAACATGTCTAGTGGGCAACAAAGTGTATGTTTTTCATGTAG
GTTCTATTACAAATGTGCTTTGAATTTATATTACATAATAGCCACTTGAACTGGAGTGAGAAGCTGGTTACTTG
CTTGCACTGCATCAGAACCAAGGTTTCTGGGGTGGCCCCCTGCAGGTTCACTCCGCCAGGAGATCAAAAAGGATG
50 CGCTCTCTACATTCTTAAATGCCACAGTCCCCACCAACGGCTTAGGTGGAAGCACAGCTCCCATCATCAGCACTG
ATGCCCTGGCGCGGCCAGCACTGCCTCTACGAGGCCACACCTGCCTCTGCCAAGGCCACTGGCCCTCATTCATCT
GCGGGGACGTTCTAGAGATATGCACTTACAG

SEQ ID NO:292

>GL3_4.03_B_M13R Direction: anti-sense

55 CTGTAAGTGCATATCTCTAGAACGTCCCCGAGATGAATGAGGGCCAGTGGCCTTGGCAGAGGCAGGTGTGGCCT
CGTAGAGGCAGTGCTGGCCGCGCCAGGGCATCAGTGCTGATGTGGGAGCTGTGCTTCCACCTAAGCCGTTGGTAG
GGGACTGTGGCATTAAAGAATGTAGAGAGCGCATCCTTTTGATCTCCTGGGCGGAGTGAACCTGCAGGGGCCAC
CCAGAAACCTTGGTTCTGATGCACTGCAAGCAAGTAACCAGCTTCTCACTCCAGTTTCAAGTGGCTATTATGTA

ATATAAATTCAAAGCACATTGTGAATAGAACCTACATGAAAACATACACTTTGTTGCCCACTGACATGTTACCAG
AAGTTGTACCATGATGTTGTTTTGACCCCTG

SEQ ID NO:293

>GL3_4.03_C_M13F Direction: anti-sense

CTGTAAAGTGCATATCTCTAGAACGTCCCCGCGAGATGAATGAGGGCCAGTGGCCTTGGCAGAGGCAGGTGTGGCCT
CGTAGAGGCAGTGTCTGGCCGCGCCAGGGCATCAGTGTCTGATGTGGGAGCTGTGCTTCCACCTAAGCCGTTGGTAG
GGGACTGTGGCATTAAAGAATGTAGAGAGCGCATCCTTTTTGATCTCCTGGGCGGAGTGAACCTGCAGGGGCCAC
CCCAGAAACCTTGGTTCTGATGCACTGCAAGCAAGTAACCAGCTTCTCACTCCAGTTTCAAGTGGCTATTATGTA
ATATAAATTCAAAGCACATTGTGAATAGAACCTACATGAAAACATACACTTTGTTGCCCACTGACATGTTACCAG
AAGTTGTACCATGATGTTGTTTTGACCCCTG

SEQ ID NO:294

>GL3_4.03_C_M13R Direction: sense

CAGGGGTCAAACAACATCATGGTACAACCTTCTGGTAACATGTCTAGTGGGCAACAAAGTGTATGTTTTCATGTAG
GTTCTATTACAAATGTGCTTTGAATTTATATTACATAATAGCCACTTGAAACTGGAGTGAGAAGCTGGTTACTTG
CTTGCACTGCATCAGAACCAAGGTTTCTGGGGTGGCCCTGCAGGTTCACTCCGCCCAGGAGATCAAAAAGGATG
CGCTCTCTACATTCTTAAATGCCACAGTCCCCTACCAACGGCTTAGGTGGAAGCACAGCTCCACATCAGCACTG
ATGCCCTGGCGCGGCCAGCACTGCCTCTACGAGGCCACACCTGCCTCTGCCAAGGCCACTGGCCCTCATTCATCT
GCGGGGACGTTCTAGAGATATGCACTTACAG

SEQ ID NO:295

>GL3_4.03_D_M13F Direction: anti-sense

CTGTAAAGTGCATATCTCTAGAACGTCCCCGCGAGATGAATGAGGGCCAGTGGCCTTGGCAGAGGCAGGTGTGGCCT
CGTAGAGGCAGTGTCTGGCCGCGCCAGGGCATCAGTGTCTGATGTGGGAGCTGTGCTTCCACCTAAGCCGTTGGTAG
GGGACTGTGGCATTAAAGAATGTAGAGAGCGCATCCTTTTTGATCTCCTGGGCGGAGTGAACCTGCAGGGGCCAC
CCCAGAAACCTTGGTTCTGATGCACTGCAAGCAAGTAACCAGCTTCTCACTCCAGTTTCAAGTGGCTATTGTGTA
ATATAAATTCAAAGCACATTGTGAATAGAACCTACATGAAAACATACACTTTGTTGCCCACTGACATGTTACCAG
AAGTTGTACCATGATGTTGTTTTGACCCCTG

SEQ ID NO:296

>GL3_4.03_D_M13R Direction: sense

CAGGGGTCAAACAACATCATGGTACAACCTTCTGGTAACATGTCTAGTGGGCAACAAAGTGTATGTTTTCATGTAG
GTTCTATTACAAATGTGCTTTGAATTTATATTACACAATAGCCACTTGAAACTGGAGTGAGAAGCTGGTTACTTG
CTTGCACTGCATCAGAACCAAGGTTTCTGGGGTGGCCCTGCAGGTTCACTCCGCCCAGGAGATCAAAAAGGATG
CGCTCTCTACATTCTTAAATGCCACAGTCCCCTACCAACGGCTTAGGTGGAAGCACAGCTCCACATCAGCACTG
ATGCCCTGGCGCGGCCAGCACTGCCTCTACGAGGCCACACCTGCCTCTGCCAAGGCCACTGGCCCTCATTCATCT
GCGGGGACGTTCTAGAGATATGCACTTACAG

SEQ ID NO:297

>gi|22051279|ref|XM_040265.5| Homo sapiens KIAA0217 protein (KIAA0217),
mRNA

AGCATCCTGAGGTTAGCATTACACTTCTCCAGTGGAGCCCATGACTTCTGATCAGGACGCTAAGGTTGT
GGCTGAACCGCAGACGCAGAGAGTCCAGGAGGGCAAGGACAGCGCTCATCTGATGAATGGTCCTATATCT
CAAACCACTTCTCAGACAAGTTCCATCCCACCTTTGAGTCAGGTACCAGCAACTAAGGTTTCAGAGCTGA
ACCCTAATGCAGAAGTGTGGGGGGCTCCTGTGTTACATCTGGAAGCAAGCAGTGTCTGACGGTGTGAG
TGCTGCATGGGAGGAGGTGGCTGGCCACCACGCAGACCGTGGCCCCGAGGGATCGGATGCCAATGGTGAT
GGTGACCAGGGCCATGAGAATGCCGCATTGCCAGACCCGAGGAGTCGGACCCAGCAGACATGAACGCTC
TCGCTCTGGGTCCCTCAGAATATGACTCTCTGCCTGAAAATAGCGAGACAGGAGGAAATGAGTCTCAACC
AGACAGCCAGGAAGACCCCCGAGAAGTACTTAAAAAACATTGGAATTCTGCTTATCTAGGGAGAACCTT
GCTAGTGACATGTATCTTATATCACAGATGGATAGTGACCAGTATGTGCCAATCAACGGTGGCTAACCC
TCGACCACATCAAGAAGCTCAGCACTGATGTGGACTTGATTGTGGAAGTGCTAAGATCTTTACCTTTAGT
CCAAGTGGATGAAAAGGGAGAAAAAGTAAGGCCAAATCAAAATCGCTGCATAGTAATATTGCGTGAAATA
TCTGAATCTACCCCGTGGAAGTAATAGAAGCACTATTTAAAGGAGATAATTTACCAAAATTTATAAAT
GTGAATTTGCATATAATGATAATTGGTTTATTACATTTGAAACAGAAAGCTGATGCACAACAGGCTTACAA
ATACCTTCGAGAAGAAGTCAAACTTTTCAAGGAAAACCAATTAAGGCACGGATAAAAGCAAAGGCAATA
GCTATAAACACATTTTGGCCAAAGAATGGATTTAGACCCCTGGACGTGAGCCTGTATGCCCAGCAGCGCT
ACGCGACGTCGTTCTACTTCCCTCCCATGTACAGCCCCCAGCAGCAGTTCCCCCTGTACAGCCTGATCAC
TCCCAGACGTGGTCAGCAACGCACAGCTATCTTGACCCACCTTGGTAACTCCATTTCCAAATACTGGA
TTTATAAATGGGTTTACGTCTCCAGCGTTCAAGCCTGCGGCGTCTCCTCTGACTTCTCTCAGACAGTATC

CTCCTCGAAGCAGGAATCCTAGTAAATCTCATCTGCGGCATGCGATTCTAGTGCAGAGAGGGGACCTGG
GTATTAGAAAAGTCCTTCAATATTTAACTTTCAGTGCAGATCGATTAATTAATGGTGTCCGGAGTCCACAA
ACAAGGCAAGCAGGTCAAAC TAGAACACGGATTCAAACCCCTCAGCATATGCCAAGAGAGAGGCTGGGC
CTGGGCGTGTGGAGCCAGGCAGTCTCGAATCCTCTCCTGGTTTAGGGAGGGGAAGGAAGAATTCCTTTGG
5 CTACCGGAAGAAAAGGGAGGAGAAGTTTACAAGCAGCCAGACACAGTCTCCAACGCCACCAAAGCCTCCG
TCGCCAAGCTTCGAGCTGGGGCTGTCCAGCTTCCCTCCATTACCTGGAGCTGCCGGCAATTTGAAGACAG
AGGACTTGTTTGAAAACAGGCTATCTAGCTTGATAATAGGACCATCCAAAGAAAGGACCCCTCAGTGCAGA
CGCAAGCGTGAACACCCCTTCTGTAGTGGTCTCCAGAGAGCCCCTCGGTGCCGGCTTCTTGTGCTGTATCA
10 GCAACGTACGAGCGATCCCCCTCCCCAGCTCATTTACCCGATGATCCCAAGGTGGCGGAGAAAACAGAGGG
AAACCCACAGTGTGGACAGACTTCCCTTCCGCCCTCACTGCGACCGCGTGTAAATCGGTGCAGGTGAACGG
AGCCGCCACGGAATTCGAAAAGCCAGCTACGCAGAGATTGTGTGAGAGAAGAGTAAAGAGCCCTCCTTCT
TCCCCATTGCAACCCCAAAAAGAACAAAAGCCAAACACTGTTGGTTGTGGGAAGGAGGAAAAAGAGCTGG
CAGAGCCCGCAGAGAGATACCGGGAGCCCCAGCCCTCAAGTCCACACCTGGAGCCCCCAGAGACCAGAG
15 GCGGCCGGCGGGGGGCGGCCCTCGCCCTCGGCCATGGGGAAGCGTCTCAGCCGAGAGCAGAGCACTCCC
CCCAAGTCTCCTCAGTGAACACCGTACGCTCGGGAGGGTTCGCAGAGCGCTGTGTAAACCACAAACGAGA
CACTCTCCCACTCAGTGCAGGGGCGAGCCGCTGGTTAGGAGCTTGCAGTGTCTGAGGCCTGTGGGATCCT
CAAGTTGGTTTTCTTCTGTGAGTTGGATTCTCCCCCTCTTGAAAAAAATCGATTTTTTCAGGATTTAATT
AATACAAACCTTATTTTAGGTTGGTGCCTTAAC TGGAGGTGATGCATAAGTCTGATTTTTTTTTTCCAAGAT
20 AGAAAAAGCATTATCTTAACAAATGGTATTTTTTATTAAGCCTCCATGTGGCTCTGAATGCAAGCTAT
ATATAGTGAGTTTTCTAAATTAAGGGAACCTGCTTTTTTTTTTTTTTTTAAAGTAAGTGGTCTGTGAAG
TGATATCTCTAGAACGTCCTCCGCGAGATGAATGAGGGCAGTGGCCCTTGGCAGAGGCAGGTGTGGCCCTCG
TAGAGGCAGTGTGGCCGCGCCAGGCATCAGTGTGATGTGGGAGCTGTGCTTCCACCTAAGCCGTTGG
TAGGGGACTGTGGCATTTAAGAAATGTAGAGAGCGCATCCTTTTTGATCTCCTGGGCGGAGTGAACCTGCA
25 GGGGCCACCCAGAAAACCTTGGTCTGTATGCACTGCAAGCAAGTAACCAGCTTCTACTCCAGTTTCAAG
TGGCTATTATGTAATATAAATTCAAAGCACATTGTGAATAGAACCTACATGAAAACATACACTTTGTTGC
CCACTGACATGTTACCAGAAAGTTGTACCATGATGTTGTTTTGACCCCTGTGAGCTGATGGCCCCGGCCCT
GCTCTGTGCACATTTCTGTCCGTGTTCCCCAGCACTCTGGTTGGAGAGAGTCCACATCTTCAGCTCCGTG
TGGAGATCTCCCTGTACCTCTGCATCAGCACTGGATTAAAGAGTTATGTAATCGTGAGAGAATGGTGT
30 TGTGGTTTTTCCCCCTCTTTGGCTGGTGGAGGATAAAGTTCCTGCTCTTTTACCTCCAAGACGAGGCCCT
CATTGATTCATTCCAGAAAGTCTGCACCTCTGAAGAACAAGGATGCACTAAAGTTAGCAAGTTTATAAT
AAAGTTAAATATAAATATTTTGTTTTAAATGCTCCTCAATTTTTCTTTATTCTAAGCAGCAAAACATTAA
AATAAGAATATTTCTGCTAAATGTAACCATAACATTTATTCCACAAAATGTTATTTAACAAGACTGAGG
GTTTTTTTTTAAGAAAAAATTTATTTCCATCCAATATTTAAAGACTTGAATTTTATTTAAACTTGAAAATGA
35 CTTTGCCCTTAACCTTTTGATATAAGACAGCTTAGAGTCCATGGAGCCCGGCCCTGGGTGGCGTGAGTGGGT
CAGAGTTACTCAGTTACTGCGTGGATCTCCTGTGCTAGTTTTACTGAGTAAGCATACTGTAGTACAAGA
GCTAGTAGTAGTTTTTGTAAATATACCTTAAAGATCTTCAACAGTTGATCTTTTTTTCAGAAATGTTGAAAA
TCCTGTAAATGCAAAATAGTCAATACTGTATTAATACAGTGCACCTGGAGTGTGCTTCGCTTGTACAGTTG
TAAATAATCAGAACATAGAAAAGGTACCTACAGAGAAAATCTGATACAGATTATGATATATATATA
40 AATGTTGCTGTTGAGCGGGATGTAGATAAACTAAATGTTGTGGTTTGAATATTATTTTGTATTGTTGAGA
TTTTCTTTTTTCTCTTACATCGGTGTGTTGAAC TGAATCTGCTCTTTGCTGCAAAAAGGGAATTGGAAAG
TCTTATTAAGCCCTCCAGATGTTTTCTACTCTTTTTAAATGTATGTAAATGCATACTAATCATATCTA
ATGTGAAAGAGTTTTTAAAGTATATAGAGAGCAAAAAGTGGCAGGATCGTAAGTGAAGGTGACTAGTAATC
TAATTTAAATCACCTGCAGCTAAGCATGATTGACCCCTGCCAGAGGAAAACATGCCTATTTGACCATTTCC
TTTAAAGCAGTTGCCATTATTCAAATACAGAGAAATAGCCACAGGGCTAGTGTTTTTCAAATGCATTTTA
45 AAGAACATGGGGATTTTTTTTTTGTAGTTGTCTAGTTCACTGACCAAAAAAAAAAAAAAAAAATCAGAAATAAT
TGATCTGTGAAACCCAACTCTCAATACTCAGAAAGCTGGGAGGCAACCTCGAGGCCTGGGCCTACGAGC
TGATCTTTCGCTACGGAAGGGCCAGGGCGCCATCAGCCATTCCCAAAACACAAGGCCTGCCCGTCCGCCA
GTGAGTCCCTGGTTTTTTAATAATGAGAAGTCTTTCCCCCAAGGTGTGAGCATTGCAGCGCAGTGTGTGT
GTGTGGTTAGAGCCAGCTTAGTCTTCACTTTGTGCGACCGAAGTGGGAGCTCAACAGCTGCATGAGGAGG
50 GCAGCGCGTGCATTAGCCAGTCGCCACTGGAGGGCTCTGCTGCCCTCCGGTCAATACACTGTAGTTACTG
CCTAGCCAGCAGCAGTCTTCTGCATCAAGAACTGAAACGTTGCTCGGAGGTGATTTTTATAGCATCCTTT
TTAATTAAAGGTGAAATACAGATTGCTATATAATGTCTGAAAAACCTGATACTACTTCAAGAGTTTCTG
CTCAGAAGAAAATGAGAGTTATCATAATAGGAAGCTGTGGCGGTCCATGCCAACTGTGCTGTGTACATA
CAGCGATGAGAGTGGCTTTTACACTTTTTTTTTTTTTTAAAGTTAACACCCCTCCTTTACCCCCAGCAGTATC
55 TCAGGTTATAGAATCAGAGATGCAGCAGTGCAAAATGGCATTTTAACTTGTAATAATCGTGTGATGATGCT
TATCATTTTGAAATAGAAGATAAAAAACCTGGTCCCGTTTACCAGACATGAATTTCAAGTGGAGTCGTC
GTTCTCTGAGAGTGAGTGTCTTGACATTTTACCCAGGCCCTCCTGTGCATGCATCACCCGGCTGTCACTG
GCGGGTGGCCGTAAACGTCCTGCGTTGCTATATTAGGATCTCTGCAGTTTCAAGCTTCAAAACAGTTTCAG
TGTATCCGGGCGACGGGTAGTGGTGGTGCATGCCTGTCTGTGTGCCCCGCTGGCGAGCTGTAGTTGCGGC
60 TTGCGTGCCCTCGCGGCCACTACAGGGCTGCAGACAATCGAGGCGAGGGCGCTGGCCGCCAGCAGCTCAC
AGCGCGGGGTCATGTGGTGCCTCTCGAGGGTTTCGTTTTTGTCTGCTTCATTAAGACTGGAATCAAG

CTTACATGTAACTATTGGTAATTTAAGTTTCCTTTTGTGTCATTCACTGTAAACTGTCTAATTTGAAA
AAAAATGTAGGTTATGAAAATAAAGATTTAGGCACTGTTT

SEQ ID NO:298

5 >gi|20543999|ref|XP_040265.4| similar to no similarities to reported gene
products [Homo sapiens]
MTSDQDAKVVAEPQTQRVQEGKDSAHLMNGPISQTTSSQTSIPPLSQVPATKVSLELNPNAEVWVGAPVLHL
EASSAADGVSAWEEVAGHHADRGPDQSDANGDGDQGHENAALPDQESDPADMNALALGPSEYDSL PEN
SETGGNESQPDSDQEDPREVLKKTLEFCLSRENLASDMYLISQMDSDQYVPITTVANLDHIKKLSTDVDLI
10 VEVLRLSLPLVQVDEKGEKVRPNQNCIVILREISESTPVEVIEALFKGDNLPKFINCEFAYNDNWFITFE
TEADAQQAYKYLREEVKTFQGKPIKARIKAKAIAINTFLPKNGFRPLDVSLYAQQRYATSFYFPPMYSPO
QQFPLYSLITPQTSATHSYLDPLVTPFPNTGFI NGFTSPAFKPAASPLTSLRQYPPRSRNPSKSHLRH
AIPSAERGPGLLESPIFNFTADRLINGVRSPQTRQAGQTRTRIQNP SAYAKREAGPGRVEPGSLESSPG
15 LGRGRKNSFGYRKREEKFTSSQTSPTPPKPPSPFELGLSSFPLPGAAGNLKTEDLFENRLSSLIIG
PSKERTLSADASVNTLPVVVSREPSVPASCAVSATYERSPSPAHL PDDPKVAEKQRETHSVDR LPSALTA
TACKSVQVNGAATELRKPSYAEICQRTSKEPPSSPLQPQKEQKPNTV GCGKEKKLAEP AERYREPPALK
STPGAPRDQRRPAGGRPSPSAMGKRLSREQSTPPKSPQ

SEQ ID NO:299

20 Novel (KIAA1583)
>GL2_89_2B04_G3F1 Direction: N/A
ACTGCAGGTGGCAGCACGGGCGCGCGCTCGGGCTCATCTGACGCCTCTGCAGCGGGTCCCGCAGGCCTGCAGGG
CGGGGAGGCCGGGACTGGGCGTCAGCGCTGAANGGGCCCAGCTGCCAGGGGGCCAGCTGCTGGAGAACCGCAGCT
CGTCCCCGGCGGCTCCTAATCACCAGCAGCTCCTGTTTCTCAAACGCAGACATCCGCCCTCTTGGGGTCAGGCCC
25 TTCCACCTGCAGTGCAGCGCGCCAGCCACTCCCGACTGGCGCTATGCCTCGATCACC GCCTCTTGCTCCCAAGTG
GACCGCAGGGTGAGACGCTCTCTTACGTGGGACCCTGGGTGTGCGCTCACTCTCTGAAGGGCCTGGAAGCTAGAT
TCCAGAGGCGTGGGCCAACTCTCCCTGGGTTTTTGGGTGAGCCCCCTCCGAGGGTGTTTCAATTTCTGAGCTCTGT
GTCATCTTAAGCTCTGAGGGTNCGNAGAAACAGACAGTGCTTAGATAGATAAAGCGGGCGGTCGAATAATCTGA
GGTCTGATACTCACTGACTGTCTGTAAGGGCGAATTACGGGGCGNTAAAATT CNAATTGCTTATAGTGAGTCTGTA
30 TTACAAATTCAGTGGCGCGGTTTTACAAANGTCGTGACTGGGAAAACCTGGCGTTACCCA ACTTAATCGCTTGA
GGAATCCCCCTTTTCGAGTGGGGTATAGGAGAAGGCGACGAATGCCTCCAACAGTGGGCAGCNAAGTACGACGGT
AAGGGTANACCTATAAAGAAANCGGACGCGGNGGAGCACGGGAAAATACCGGAGAGAGTGNGACCCCTGCCAGAC
GGATTG

35 SEQ ID NO:300

>GL2_202_3_M13F Direction: anti-sense
ACCTCAGAGCCTAAGATGACACAGAGCTCAGGAAATGAACACCCTCGGAGGGGGCTCCCCAAAACCCAGGGGAG
AGGTGGCCACGCCTCTGGAATCTAGCTTCCAGGCCCTTTTTCAGAGAGTGAGCGCCCCCAGGGGTCCCCGTAAG
AGAGCGTCTCCCCCTGCGGTCCACTTGGGAGCAAGAGCGGTGATCGAGGCATAGCGCCAGTCGGGAGTGGGCTGGG
40 GCGGCTCGCTGCAGGGTGGAAAGGGCCTGACCCCAAGAGGGGCGGATGTCTGCGTTTTGAGAAACAGGAGCTGCT
GGTGATTAGGAGCCCGGGGACGAGCTGCGGGTCT

SEQ ID NO:301

>GL2_202_3_M13R Direction: sense
45 ACTGCAGGTGGCAGCCACGGGGCGGCGCGCTCGGCCTCATCTGACGCCTCTGCAGCGGGTCCCGCAGCCT
GCAGGGCGGGGAGGCCGGGACTGGCCGTGAGCGCTGAACGGGCCAGCCTGCCAGGCCAGCTGCTGGAGACCCG
CAGCTCTGTCCCCGGCGGCTCCTAATCACCAGCAGCTCCTGTTTCTCAAACGCAGACATTGCCCCCTCTTGGGGT
CAGGCCCTTCCACCTGCAGGCGAGCCGCCAGCCACTCCCGACTGGCGCTATGCTC

50 SEQ ID NO:302

>GL2_89_2B01_G3F1 Direction: sense
ACTGNTAGNGTGGCAGCACGGGCGCGCGCTCGGCCTCATCTGACGCCTCTGCAGCGGGTCCCGCAGNGCCTGCA
GTGGGCGGGGAGTGCCGGGACTGGCCGTGAGCGCTGAACGGGCCAGCTGGCCAGGGGCCAGCTGCTGGAGACCC
GCAGCTCGTCCCCGGCGGNTCCTAATCACCAGCAGCTCCTGTTTCTCAAACGCAGACATTCTGGCCCTCTTGGGG
55 TCAGTGCCTTCCACCTGCAGTGCAGGCTGGCCAGACCACTCCCGACTGGCGCTATGCCTCGATCACC GGTC
TTGCTCCCAAGTGGACCGCAGGGTGAGACGCTCTCTTACGGGGACCCTGGGGGCGCTCACTCTCTGAAGTGGCCT
GGAAGCTATGATTCCAGAGGCGTGGGCCAACTCTCCCTGGGTTTTGGGGAGCCCCCTCCGAGGGTGTTTCAATATC
CTGAGCTCTGGTGTCTATCTTAGGCTCTGAGGGTCTGAGAAACCAGGCACCAAGTGGTTAGATAGATAAAGCTGG
GCGGTGACATAGATCTGAGGTCTGATACTCACTGACTGTCTGTAAGGGCGAATTC

SEQ ID NO:303

>GL2_89_2B02_G3F1 Direction: sense

ACTGCAGGTGGCAGCCACGGGCGCGCCGCTCGGCCTCATCTGACGCCTCTGCAGCGGGTTCCGCAGGCCTGCAGG
GCGGGGAGGCCGGGACTGGCCGTCAGCGCTGAACGGGGCCAGCCTGCCAGGGGCCAGCTGCTGGAGACCCGAG
5 CTCTGTCCTCCCGCGGCTCCTAATCACCAGCAGCTCCTGTTTCTCAAACGCAGACATCCGCCCTCTTGGGGTCAGG
CCCTTCCACCTGCAGNGCGAGCCGCCCCAGCCCACTCCCGACTGGCGCTATGCCTCGATCACCGCTCTTGCTCCC
AAGTGGACCGCAGGGGAGACGCTCTCTTACGGGGACCTTGGGGGCGCTCACTCTCTGAAGGGCCTGGAAGCTAGA
TTCCAGAGGCGTGGGCCACCTCTCCCTGGGTTTTGGGGAGCCCCCTCCGAGGGTGTTTATTCTCTGAGCTC

SEQ ID NO:304

>GL2_89_2B03_G3F1 Direction: sense

CAGTGGTTAGATAGATAAAGCGGCCGCTCGACTAGTCTGAGGTCTGATACTCACTGACTGTGTAAGGGCGAATT
CGCGGCCCGCTAAATTCATTCGCCCTATAGTGAGTCTGATTACAATCACTGGCCGTCGTTACAACGTCGTGACT
GGGAAAAACCTGGCGTTACCCAACTTAATCGCCTTGACAGCACATCCCCCTTTCGAGCTGGCGTAATAGCGAAGA
15 GGCCCCGACCGATCGCCTCCCAACAGTGCAGCAGCTATACGTACGGGAGTAAGGGTNCACCTATNAAGAAANAGCGAC
CGCCGGTGGGCGNCACACAGGGACAANAATGGACGCACGANCAACGAACGTCAACCCAGAACAGAACTAGTCA
AAAAANCCGAAAAACNGAGAGACTAAGGAAAANNCAAAACAAAAACACAAGAAAACCAACC

SEQ ID NO:305

>GL2_89_2B05_G3F1 Direction: N/A

ACTGCAGGTGGCAGCCACGGGCGCGCCGCTCGGCCTCATCTGACGCCTCTGCAGCGGGTTCCGCAGGCCTGCAGG
GCGGGGAGGCCGGGACTGGCCGTCAGCGCTGAACGGGCCAGCCTGCCAGGGGCCAGCTGCTGGAGACCCGCA
GCTCGTCCCCGCGGCTCCTAATCACCAGCAGCTCCTGTTTCTCAAACGCAGACATCCGCCCTCTTGGGGTCAG
GCCCTTCCACCTGCAGGCGAGCCGCCAGCCCACTCCCGACTGGCGCTATGCCTCGATCACCGCTCTTGCTCCC
25 AAGTGGACCGCAGGGGAGACGCTCTCTTACGGGGACCTTGGGGGCGCTCACTCTCTGAAGGGCCTGGAAGCTAGA
TTCCAGAGGCGTGGGCCACCTCTCCCTGGGTTTTGGGGAGCCCCCTCCGAGGGTGTTTATTCTCTGAGCTCTGTG
TCATCTTAGGCTCTGAGGGT

SEQ ID NO:306

>GL2_89_2B06_G3F1 Direction: N/A

ACTGCANGTNGNAGCCACGGGCGCGCCGCTCGGGCTCATCTGACGCCTCTTTTCAGCGGGGTTCGCGAGTGCCTG
CAGGGGCGGGGAGTGCCGGGACTGGGCGTCAGCGCTGAAGGGGCCAGCTGTCCAGTGGGGCCAGCTGCTGGAGAA
ACCGCGAGGNGTCTGGTTCCCGGCGGGGTCCCTAAGTNACCAGCAGAGNGTCTGTTTCTCAAACGCAGACAT
TCTGGCCCTCTTGGGGGTGAGGGCCTTTCACCTGCAGGCGAGCCCGCGCCAGNGGCCGACGGTCCCCGAAACAT
35 GGGGCGGCGTAATTGTCTTCTCGAATCTAAACAGCGGTCC

SEQ ID NO:307

clone 89 -Novel (KIAA1583)>

ACTGCAGGTGGCAGCCACGGGCGCGCCGCTCGGCCTCATCTGACGCCTCTGCAGCGGGTTCCGCAGGCCTGCAGG
40 GCGGGGAGGCCGGGACTGGCCGTCAGCGCTGAACGGGCCAGCCTGCCAGGGGCCAGCTGCTGGAGACCCGAGC
TCGTCCCCGCGGCTCCTAATCACCAGCAGCTCCTGTTTCTCAAACGCAGACATCCGCCCTCTTGGGGTCAGG
CCCTTCCACCTGCAGGCGAGCCGCCAGCCCACTCCCGACTGGCGCTATGCCTCGATCACCGCTCTTGCTCCC
AAGTGGACCGCAGGGGAGACGCTCTCTTACGGGGACCTTGGGGGCGCTCACTCTCTGAAGGGCCTGGAAGCTAGA
TTCCAGAGGCGTGGGCCACCTCTCCCTGGGTTTTGGGGAGCCCCCTCCGAGGGTGTTTATTCTCTGAGCTCTGTG
45 TCATCTTAGGCTCTGAGGGT

SEQ ID NO:308

>gi|10047240|dbj|AB046803.1| Homo sapiens mRNA for KIAA1583 protein,
partial cds

GTGGTGAGAGGGGAAGGGACATTTACAGCCCTTTTCTCCCTCCCTTGCCACAACTACTGTGCCCCAGCA
50 TGGCACCCGGGTCAGCCCTGCCAGCTCTGGAGCTGTATCATCTCTCCCTCTTATGCCTCTTCAGTGGA
CTGTGGCCAGGCTCCCTTGGACCCTGTCTACCTGCCGGCAGCCCTGGAGCTCCTAGACGCCCCGTAACAC
TTCCGTGTGCAGCAGGTGGGCCACTACCCACCTGCCAACTCCTCTCTGAGCTCCCGATCTGAGACCTTTC
TGCTCCTACAGCCCTGGCCCAGGGCCAGCCACTTCTCCGGGCCCTCCTACCCACCTTTTGCCACTCAGCA
55 GGTGGTCCCCCTCGAGTCACTGAGCCCCACCAACGGCCAGTCCCATGGGACGTGCGGGCCGTTTTCAGTG
GAAGCGGCTGTGACTCCAGCAGAGCCCTACGCCCCGGTTCTCTTCCACCTCAAAGGGCAGGATTGGCCAC
CAGGGTCTGGCAGCCTGCCCTGTGCCCGGCTCCATGCCACACACCTGCCGGCACTGCTCACCAAGCCTG
CCGCTTCCAGCCATCCCTGGGCGCCTGCGTGGTGGAGCTGGAGCTTCCCTCGCACTGGTTCTCACAGGCC
TCCACCACACGGGCCGAGCTGGCCTACACGCTTGAGCCTGCAGCTGAGGGCCCTGGGGGCTGTGGCTCCG
60 GCGAGGAGAACGACCCTGGGGAGCAGGCCCTCCAGTGGGGGGTGTGGAGCTGCGCCAGCAGACCCCC

GCAGTACCAGGAGGTACCTCTGGACGAGGCTGTGACTCTGCGGGTGCCCTGACATGCCAGTGCGGCCCCGGC
CAGCTCTTTAGTGCTACCCCTCCTGCTTCGGGCACAACCTTACAGCCAGCCCTCCTGACCCTGCGGATCAAGG
TGAAGAAGGGGCTGCATGTGACAGCCGCCCCGCCCAGCCCAGCCCACACTCTGGACTGCCAAGCTGGACCG
TTTCAAGGGCTCCAGGCACCACACCACCTCATCACCTGCCACCGTGTGGGCTCACAGAGCCAGATTCC
5 AGTCCCCCTTGAACCTGTCTGAGTTCCTATGGGTGGACTTTGTGGTGGAGAATAGCACTGGTGGGGGCGTAG
CGGTCACTCGCCCCGTACGTGGCAGCTGGAGTACCCAGGCCAGGCCCTGAAGCAGAGAAGGACAAAAT
GGTGTGGGAAATCCTGGTGTCTGAGCGGGACATCAGAGCCCTTATCCCACTGGCCAAGGTAAGGAGACCT
CCATCTCTGCCTGGGAAGGCTGGAGGCCAAGTCCCAGGAGCCCCATGAGCTGAGGGGCTGGGGCTGAGC
10 CCCTCCTCCACCCCCAGGCTGAGGAGCTGGTGAATACAGCACCCTGACTGGAGTGCCCCAGCATGTCCC
CGTGCCTTGTCACTGTGGACGGCGGGGGGCTTGGTGGAGGTGACAGAGCATGTGCGCTGCGAGTCT
GCCAACACACAGGTCTCTGAGGTGAGTGGCAGGTGCCAGCTCATGTGAGTCTGCATTTGTGTGGGCACA
GTTACGCACACACACCTTTCTCTCTCTCATGTGGGTCCCCAAGAGAACAGCTTCTTCTTGAGACGAGA
AACCTCATGGGAGGAACCACAGCTTCAACTTCTTTTGTTCAGCCACCTCTCTCCCGAGTCAAGTAACCC
TAACAGCCCCCGACCTTAGACACTTTCACGCTCATTATTGCACCTGGCAAGTGCCTGGCCTGAGAGAGT
15 CGAGAGCACATACTCTGGGATCAGAAAGTTTGTAGTTTGAACACATCTTTGTGTCAACCAGCTTTGTGA
TCATGGGCTGGTCTCTTAACCTCTCTGAGCCTCGTTACCTAAACTTTAAATGAGAACATATCTTCGGGT
TGTTGTAAGAATTAAGCACTTTAAGACAGTATTTAGCAGATAGCAGGGGCTGAAACAATGGCAGTTAT
TGCTGTTATAATATGATGTAATAATTATACAATGGTAAAAATTATTATTATTATTATTACCGGTGATTGT
GAAGTGAAATAAAGGGAGGGTGGTAGAGCAACTTCTCTATCTTGAAGTTCTAAGAGGGGAAGCATGGA
20 GTCTGCATGTATGCAAATCCAGCCACCTGAAAGGGGAAGGTGCCTTTCCCACTTGTGGTCTTTGTCTC
AAATGTGTGTGGTTCTGTCTGGAGGGCTTGGTGGGAGATCTGCTTTGTCCCACTCTGCCTACCTGGCT
GTAGCTGGTTGCCAGGTGGCTCCAAGAGCAATGTCAAAGGCCAGCATGAATAAGGAGCCAAAGTTTGGG
TTCACACCCCTGCTCTGCCAGGAGGAAGACTGGTATGCAATGGCCAGTTCCCTGACTGAGCCCCAGGGA
ACCTCAAGGATAGCTCCCTACAGTAAATCTTAAACAAAGAACAGGGGCTTAACCAACAATCTGTCTGCA
25 ACATGCATTCCAGTGACACACAGAGCTGAACCTCAGTGGGAGCCCGTCTACGGCCTGAGGTGAGGGCTGC
TCTCTTCTGGCCCATGCCTTTTTTTTTTTTTTAATCTATTTATTTTGAACAGAGTCTCACTCTGTCCCC
AGGCTGGAGTGCAATGGTACAGTCCGGCTCACTGCAGTCTCAGCTACCCAGGCTCAAGGGATCCTCCAC
CTCAGCCTCCCACTAGCTAGGACTACAGGCATGTGCCACCACACCCGGCTGATTTATTGTATTTTTTGT
AGAGAGGGAGTTTTTGCCATGTTGCCAGGCTGGTCTCAAACCTCCTGGACTCAAGCGATCCACCCACCTCG
30 GCCTCTCAGAGTGCTGGGACTACAGGCATGAGCCACCCACCCGGCCAGCCTCACACCTGTCTTTAGTGG
CTTAGCCCGCCACCTGCCCTCTGCCCTTCCATAGGTGTCTGAGGCCTGTGATGCCGTGTTCTGTGGCTGG
CAAGGAGAGCCGGGCGCCCGGGGGTGCAGTGGACTTCTGGTGGCGCCGGCTCCGCGCCTCGCTGCGG
CTGACCGTGTGGGCCCCCTGCTACCGCTGCGTATCAGCTACCGACACCACCTCGAGCAGGTCCGCG
GCTGGAGGGTACCTGGCCCTGCTGAAGGGTGAAGTGGAGGCCTGAGGAAGCTGGCAGGCCTTGGCGAGTCA
35 CACTGGGACGGAGAGGGGCGCAGGGGACACAGCCATGGGCCAAGTCGGGGTCAGAAAACTGCCCTGTTG
CTTGCAAGTTCCTTCTGCCACCTCCACCCTGGGCCGCCAGCCAATGATCGCTTTCTGTGAAATTGCTGA
CTTCCGCATCCCTCTCCTCTTCCCTGCCTCACTTCTCCCGTAGCACCTCCACTTCCCTAATCTACGTCA
GTACTGGTTAATGTTGCCTGTCCATCCTTTCCACGCTCCTCCCCGAGAATGTGAGCCCTTGAAGCTGC
GGTCTGTCTATGCTGGTTCGTACCTCAGAGCCTAAGATGACACAGAGCTCAGGAAATGAACACCCCTCGGA
40 GGGGGCTCCCCAAAACCCAGGGAGAGGTGGCCACGCTCTGGAATCTAGCTTCCAGGCCCTTCAGAGAG
TGAGCGCCCCAGGGTCCCCGTAAGAGAGCGTCTCCCTGCGGTCCACTTGGGAGCAAGAGCGGTGATCG
AGGCACAGCGCAGTCGGGAGTGGGCTGGGGCGGCTGCGCTGCAGGTGGAAGGGCTGACCCCAAGAGGG
GCGGATGTCTGCGTTTGAAGAACAGGAGCTGCTGGTGAATTAGGAGCCGCCGGGGACGAGCTGCGGGTCTC
CAGCAGCTGGGCCCTGGGCAGGCTGGGGCGCTCAGCGCTGACGGCCAGTCCCGGCCTCCCCGCCCTGCA
45 GGCCTGCGGAACCCGCTGCAGAGGCGTGGATGAGGCCGAGCGGCGCCCGTGGCTGCCACCTGCAGTA
CCAGCGGGCCGTTGTGCGCTTCTCGCCCCCTTCGCGGCCACCCGCTGGACGGCGGCCG

SEQ ID NO:309

>gi|10047241|dbj|BAB13409.1| KIAA1583 protein [Homo sapiens]
50 VVRGEGTFTALFSLPCPQLLCPSMAPGSAPASSGAVIISPSYASSVDCGQAPLDPVYLPAALELLDAPEH
FRVQVGHYPANSSLSRSETFLLLQWPRAQPLLRSYPPFATQQVVPVRVTEPHQRPVPWDVRAVS
EAAVTPAEPYARVLFHLKGGDWPPGSGSLPCARLHATHPAGTAHQACRFQPSLGACVVELELP SHWFSQA
STTRAEALAYTLEPAEPPGGSGSEENDPGEQALPVGVELRPADPPQYQEVPLDEAVTLRVDPMPVRPG
QLFSATLLLRHNFTASLLTLRIKVKKGLHVTAARPAQPTLWTAKLDRFKGSRHHTTLITCHRAGLTEPDS
55 SPLELSEFLWVDFVVENSTGGGVAVTRPVTWQLEYPGQAPAEKDKMVWEILVSEDIRALIPLAKVRRP
PSLPKGAGGQVPGAP

SEQ ID NO:310

Novel (KIAA1814)

>GL2_75_2_M13F Direction: sense

CTGGACGGCCTGGCTGGGCTGAAGGGCGAGGACAGCCGAGCAAGGAGGCAGGGGAGGGCGGCCTACCGCTGTGC
5 GGGCCACGGACAAGACCCCACTGCTGAGCGGCAAGGCCGCCAAGGCCCGGACCGCGAG

SEQ ID NO:311

>GL2_75_2_M13R Direction: anti-sense

CTGCCCATGGCGTGCAGATTGGTCAGTCTGGTGGCAGCGGACGACACGTGCACCATGGCCTGCGGGCGGCCGGCAC
10 AGAGCTGAACAGCGACTGCAGCACGGAGGAGCCGGCCACGGAGCCGAGGTTGGCCTGCAGGGACATGGGGCCGAG
CGCGAACTGCGGTCCCGGCGGGAAGGGGCCAGGAAGGACCGGTGCGTCTGGGCGGAGGACGCCGCGCCGCCTGC
GGAGGACGCCGCCGGGGCCAGGCCGGGGCCGCTGAGGAGGCTTCCGGGAGGCACGGCCGCCGAGAGATGAAGAG
GTTGTGGCCATTCTTGAGGTCGACCTCGCGGTCCCGGGCCTTGGCGGCCTTGCCGCTCAGCAGTGGGGTCTTGTC
15 CGTGGGCGCCGACAGNGGTAGGCCGCCTCCCTGCCTCCTTGCTGCGGCTGTCTCGCCCTCAGCCAGCCAGG
CCGTCCAG

SEQ ID NO:312

>GL2_75_3_M13R Direction: anti-sense

CTGCCCATGGCGTGCAGATTGGTCAGTCTGGTGGCAGCGGACGACACGTGCACCAAGGCCTGCGGGGCGGCACAGA
20 GCTGAACAGCGACTGCAGCACGGAGGAGCCGGCCACGGAGCCGAGGTTGGCCTGCAGGGGACATGGGGCCGAGCG
CGAACTGCGGTCCCGGCGGGAAGGGGCCAGGAAGGACCGGTGCGTCTGGGCGGAGGACGCCGCGCCGCCTGCGG
AGGACGCCGCCGGGGCCAGGCCGGGGCCGCTGAGGAGGCTTCCGGGAGGCACGGCGCCGAGAGATGAAGAGGTT
GTGGCCATTCTTGAGGTCGACT

SEQ ID NO:313

>GL2_81A_D07_G3F1 Direction: N/A

CTGGACGGCCTGGCTGGGCTGAAGGGCGAGGACAGCCGAGCAAGGAGGCAGGGGAGGGCGGCCTACCGCTGTGC
GGGCCACGGACAAGACCCCACTGCTGAGCGGCAAGGCCGCCAAGGCCCGGGACCGCGAGGTCGACCTCAAGAAT
30 GGCCACAACCTCTTTCATCTCTGCGGCGGCCGTGCCTCCCGGAAGCCTCCTCAGCGGCCCGGCCTGGCCCCGCG
GCGTCTCCGAGGCGGCGCGGCGTCTCCGCCAGACGCACCGGTCTTCTGGGCCCTTCCCGCCGGGACCG
CAGTTCGCGCTCGGCCCATGTCCCTGCAGGCCAACCTCGGCTCCGTGGCCGGCTCCTCCGTGCTGCAGTCGCTG
TTCAGCTCTGTGCCGCGCCGCGAGGCCTGGTGCACGTGTCGTC

SEQ ID NO:314

>GL2_81A_D08_G3F1 Direction: N/A

CTGGACGGCCTGGCTGGGCTGAAGGGCGAGGACAGCCGAGCAAGGAGGCAGGGGAGGGCGGCTACCGCTGTGCG
GGCCCACGGACAAGACCCCACTGCTGAGCGGCAAGGCCGCCAAGGCCCGGGACCGCGAGGTCGACCTCAAGAATG
GNCACAACCTCTTTCATCTCTGCGGCGGCGTGCCTCCCGGAAGCCTCCTCAGCGGCCCGGCCTGGCCCCGCGGCG
TCCTCCGAGGCGGCGGCGTCTCCGCCAGACGCACCGGTCTTCTGGGCCCTTCCCGCCGGGACCGAG
40 TTCGCGCTCGGCCCATGTCCCTGCAGGCCAACCTCGGCTCCGTGGCCGGCTCCTCCGTGCTGCAGTCGCTGTTCA
AGCTCTGTGCCTGGCCTGCCGAGGCCTGGTGCACGTGTCGTCGCTGCCANACAGACTGACCAACTCGCACGCC
ATGGGCAG

SEQ ID NO:315

>GL2_81A_D11_G3F1 Direction: N/A

CTGGACGGCCTGGCTGGGCTGAAGGGCGAGGACAGCGCAGCAAGGAGGCAGGGGAGGGCGGCTACCGCTGTGCGG
CCCACGGGACAAGACCCCACTGCTGAGCGGCAAGGCCGCCAAGGCCCGGGACCGCGAGGTCGACCTCAAGAATGG
GCACAACCTCTTTCATCTCTGCGGCGGCGTGCCTCCCGGAAGCCTCCTCAGCGGCCCGGCCTGGCCCCGCGTGGC
TGTCCTCCGAGTGCTGGCTGCGGCGTCTCCGCCAGACGCACCGGTCTTCTGGGCCCTTCCCGCCGGGAC
50 CGCAGTTCGCGCTCGGCCCATGTCCCTGCAGGCCAACCTCGGCTCCGTGGCCGGCTCCTCCGTGCTGCAGTCGCTG
TTCAGCTCTGTGCCGCGCCGCGAGGCCTGGTGCACGTGTCGTCGCTGCCACAGACTGACCAACTCGCACGCC
ATGGGCAG

SEQ ID NO:316

>GL2_81A_D12_G3F1 Direction: sense

CTGGACGGCCTGGCTGGGCTGAAGGGCGAGGACAGCCGAGCAAGGAGGCAGGGGAGGGCGGCTACCGCTGTGC
GGGCCACGGACAAGACCCCACTGCTGAGCGGCAAGGCCGCCAAGGCCCGGGACCGCGAGGTCGACCTCAAGAAT
GGGCACAACCTCTTTCATCTCTGCGGCGGCGTGCCTCCCGGAAGCCTCCTCAGCGGCCCGGCCTGGCCCCGCGGCG
GTCCTCCGAGGCGGCGGCGTCTCCGCCAGACGCACCGGTCTTCTGGGCCCTTCCCGCCGGGACCGCA
60 GTTCGCGCTCGGCCCATGTCCCTGCAGGCCAACCTCGGCTCCGTGGCCGGCTCCTCCGTGCTGCAGTCGCTGTTT

AGCTCTGTGCCGGCCGCCGAGGCTGGTGCACGTGTCGTCCGCTGCCACCAGACTGACCAACTCGCACGCCATG
GGCAG

SEQ ID NO:317

5 Clone 75 -Novel (KIAA1814)>
CTGGACGGCCTGGCTGGGCTGAAGGGCGAGGACAGCCGAGCAAGGAGGCAGGGGAGGGCGGCCTACCGCTGTGC
GGGCCCCACGACAAGACCCCACTGCTGAGCGGCAAGGCCGCAAGGCCGGGACCGCGAGGTCGACCTCAAGAAT
GGCCACAACCTCTTCATCTCTGCGGCGGCCGTGCCTCCCAGGAAGCCTCCTCAGCGGCCCGGCCTGGCCCCGGCGG
10 CGTCCTCCGCGAGGCGGCGCGGCGTCTCCGCCAGACGCACCGGTCTTCTTGGGCCCTTCCCGCCGGGACCGC
AGTTCGCGCTCGGCCCATGTCCCTGCAGGCCAACCTCGGCTCCGTGGCCGGCTCTCCGTGCTGCAGTCGTGT
TCAGCTCTGTGCCGGCCGCCGAGGCTGGTGCACGTGTCGTCCGCTGCCACCAGACTGACCAACTCGCACGCCA
TGGGCAG

SEQ ID NO:318

15 >gi|22053031|ref|XM_046822.3| Homo sapiens KIAA1814 protein (KIAA1814),
mRNA
GCGGCCGTCCACTGGACTCCTGCGCCATATCCTGCAGCAGGTCTACAACCACTCGGTGACCGACCCCGAG
AAGCTCAACAACTACGAGCCCTTCTCCCCGAGGTGTACGGGGAGACCTCCTTCGACCTGGTGGCCAGAG
TGATTGATGAGATCAAGATGACCGACGACGACCTGTTTGTGGACTTGGGGAGCGGTGTGGGCCAGGTCTGT
20 GCTCCAGGTTGCTGCTGCCACCAACTGCAACATCACTATGGCGTCGAGAAAGCAGACATCCCGGCCAAG
TATGCGGAGACCATGGACCGCGAGTTCAGGAAGTGGATGAAATGGTATGGAAAAAGCATGCAGAATACA
CATTGGAGAGAGGCGATTTCTCTCAGAAGAGTGGAGGGAGCGAATCGCCAACACGAGTGTTATATTTGT
GAATAATTTTGCCTTTTGGTCTGAGGTGGATCACCAGCTGAAGGAGCGGTTTGCAAACATGAAGGAAGGT
GGCAGAAATCGTGTCTCTGAAACCCCTTTGCACCTCTGAACCTCAGAATAAACAGTAGAAACTTGAGTGACA
25 TCGGCACCATCATGCGCGTGGTGGAGCTCTCGCCCCCTGAAGGGCTCGGTGTCTGGACGGGGAAGCCAGT
CTCCTACTACCTGCACACTATCGACCGCACCATACTTGAAAACTATTTTCTAGTCTGAAAAACCCAAAA
CTCAGGAGGAACAGGAGGAGCCCGGCCGCCGACGAGCGCGAGCAAGAGCAACGCGGCCACGCCCCA
CTAAGGGCCCCAGGGCAAGGTGGCCGGCCCCGCGACGCCCCCATGGACTCTGGTGCTGAGGAAGAGAA
GGCGGGAGCAGCCACCGTGAAGAAGCCGTCTCCCTCCAAAGCCCGCAAGAAGAAGCTAAACAAGAAGGGG
30 AGGAAGATGGCTGGCCGCAAGCGCGGGCGCCCCAAGAAGATGAACACTGCGAACCCCGAGCGGAAGCCCA
AGAAGAACCAAACTGCAGTGGATGCCCTGCACGCTCAGACCGTGTCTCAGACGGCGGCCTCTCACCCCA
GGATGCCCTACAGATCCCCCTCACAGCCCGTTCTACCAGCTACCTCCGAGCGTGCAGCGGCACCTCCCCAAC
CCGCTGCTGGTGGCGCCACCCCGCCCCGCGCTGCAGAAGCTTCTAGAGTCCTTCAAGATCCAGTACCTGC
AGTTCCTGGCATAACAAAGACCCCCCAGTACAAGGCCAGCCTGCAGGAGCTGCTGGGCCAGGAGAAGGA
35 GAAGAACGCCCAGCTCCTGGGTGCGGCTCAGCAGCTCCTCAGCCACTGCCAGGCCCAGAAGGAGGAGATC
AGGAGGCTGTTTCAGCAAAAATTGGATGAGCTGGGTGTGAAGGCGCTGACCTACAACGACCTGATTCAAG
CGCAGAAGGAGATCTCCGCCATAACCAGCAGCTGCGGGAGCAGTCCGAGCAGCTGGAGCAGGACAACCG
CGCGCTCCGCGGCCAGAGCTTGACGCTGCTCAAGGCTCGCTGCGAGGAGCTGCAGCTGGACTGGGCCACG
CTGTGCTGAGAGAAGCTGTTGAAGGAGAAGCAGGCCCTGAAGAGCCAGATCTCGGAGAAGCAGAGGCACT
40 GCCTGGAGCTGCAGATCAGCATTGTGGAGCTAGAGAAGAGCCAGCGGCAGCAGGAGCTCCTGCAGCTCAA
GTCCTGTGTGCCGCTGACGACGCCCTGTCCCTGCACCTGCGTGGGAAGGGCGCCCTGGGCCGCGAGCTG
GAGCCTGACGCCAGCCGGTGCACCTGGAGCTGGACTGCACCAAGTTCTCGCTGCCTCACTTGAGCAGCA
TGAGCCCGGAGCTCTCCATGAACGGCCAGGCTGCTGGCTATGAGCTCTGCGGTGTGCTGAGCCGGCCTTC
GTCGAAGCAGAACACGCCCCAGTACCTGGCCTACCCCTGGACCAGGAGGTGGTGGCCTGTACCCCTAGC
45 CACGTCGGCCCGGCCGCGCTGGAGAAGCTGTCTGGCCTAGCCGCACCCGACTACACTAGGCTGTCCCCGG
CCAAGATTGTGCTGAGGCGGCACCTGAGCCAGGACCACACGGTGCCCGGCAGGCCGCTGCCAGTGAGCT
GCATTGAGAGCTGAGCACACCAAGGAGAACGGCCTTCCCTACCAGAGCCCCAGCGTGCTGGCAGCATG
AAGCTGAGCCCTCAGGACCCGCGGCCCTGTCCCTGGGGCCTTGACGCTTGTGAGAGAGAAGAGCAGTG
AGAAGGGCCTGAGAGAGCGCGCTACGGCAGCAGCGGGGAGCTCATACAGCCTGCCCATCAGCATCCC
50 GCTCAGCACCGTGCAGCCCAACAAGCTCCCGGTGAGCATTCCCCTGGCCAGCGTGGTGTGCCAGCCGC
GCCGAGAGGGCGAGGAGCACCCCAAGTCCCGTGTGTCAGCCCCGTGACCCCTCGTCCACACTTGAAAAGC
AGATTGGTGTAAATGCCACGGTGTGAGGAGCAGAAGCCTTGCCCTGGCCCCCGCAGGCTTCTCCTACGC
TGGCTCGGTGGCCATCAGCGGGGCTTGGCGGGCAGCCCGGCCCTCTCTCACACCTGGAGCCGAGCCGGCC
ACCTTGGATGAGTCTCCAGCTCTGGGAGCCTTTTGGCACCGTGGGGTCCCGCAGCTCCACGCCACAGC
55 ACCCCCTGCTGCTGTCACAGCCCCGGAACTCGCTTCTGCTCTCCCCGCCACCAGCTCTCCTCCAGTCC
CCGGCTTGGTGGGGCCGCCAGGGCCGTTGCCCGAGGCCAGCAAGGGAGACCTGCCCTCCGATTCCGGC
TTCTCAGATCCTGAGAGTGAAGCCAAGAGGAGGATTGTGTTACCATCACCACTGGTGGGGCAGTGCCA
AGCAGTCGCCCTCCAGCAAGCACAGCCCCCTGACCGCCAGCGCCCGTGGGGACTGTGTGCCGAGCCACGG
GCAGGACAGTCGCAGGCGCGGCCGGCGGAAGCGAGCATCTGCGGGGACGCCAGCTTGAGCGCAGGCGTG
60 TCCCCCAAGCGCCGAGCCCTGCCGTCCGTGCTGGCCTTTTACACAGCCTTCGGGGTCTCCCTCAACC
TCAACTCCATGGTCAGTAACATCAACCAGCCCCCTGGAGATTACAGCCATCTCGTCCCCGGAGACCTCCCT

GAAGAGCTCCCCTGTGCCCTACCAGGACCACGACCAGCCCCCGTGCTCAAGAAGGAGCGGCCTCTGAGC
CAGACCAATGGGGCACACTACTCCCCACTCACCTCAGACGAGGAGCCAGGCTCTGAGGACGAGCCAGCA
TGGTGGTGGCTTGGCGGGAAGGAGCCCGCGCCCGCGGCGAGCCAGTCAATAGCAGCAAGTGGAAGTCC
5 ACCTTCTCGCCCATCTCCGACATCGGCCTGGCCAAGTCGGCGGACAGCCCGCTGCAGGCCAGCTCCGCCC
TCAGCCAGAATCCCCTGTTTACGTTCCGGCCCGCCCTGGAGGAGCCCTCTGCCGATGCCAAGCTGGCCGC
TACCCCAGGAAAGGCTTTCCCGGCTCCCTGTGCGGGGCTGACGGAATCAGCCCGGGCACCAACCTGCC
AACGGCTGCACCTTCGGCGGGGGCCTGGCCGCGGACCTGAGTTTACACAGCTTCAGTGATGGTGCTTCTC
10 TTCCCCACAAGGGCCCCGAGGCGGGCCGGCCTGAGCTCCCCGCTGAGCTTCCCCCTCGCAGCGCGGCAAGGA
GGGCTCGGACGCCAACCCTTTTCTGAGCAAGAGGCAGCTGGACGGCCTGGCTGGGCTGAAGGGCGAGGGC
AGCCGCGGCAAGGAGGCAGGGGAGGGCGGCCTACCGCTGTGCGGGCCACGGACAAGACCCCACTGCTGA
GCGGCAAGGCCGCAAGGCCCGGAGCGGAGGTGCAGCTCAAGAATGGCCACAACCTCTTCATCTCTG
GGCGGCGGTGCTTCCCGGAAGCCTCCTCAGCGGCCCGCGCCTGGCCCCGGCGGCGTCTCCGCGAGGCGGC
15 GCGGCGTCTCCGCCCAGACGCACCGGTCTTCTGGGCCCTTCCCGCCGGACCGCAGTTCCGCGCTCG
GCCCCATGTCCCTGCAGGCCAACCTCGGCTCCGTGGCCGGCTCCTCCGTGCTGCAGTCGCTGTTTACGCTC
TGTGCCGGCCCGCGCAGGCCTGGTGCACGTGTCTCGCTGCCACCAGACTGACCAACTCGCACGCCATG
GGCAGCTTTTCCGGGTGGCAGGCGGCACAGTTGGAGGTGTCTTTAACCACGCGGTGCCCTCCGCTCTG
CTCATCCGTTTGGAGCCCGTGTGCGCCGCGGGGCTGCATGTGGCAGCGCCACACTGGGCCCCGAGCCCGCT
20 GCAGGCGGCGGCCAGCGCCTCGGCCTCTTCTTTTCCAGGCCCGGCCCTCGGTTGAGACCCGGCCGCCCT
CCGCTCCGCTCCGCTCCCCCGCTGCCCCCGCTGCGCACCTGGGCGGTCGCCCGGGGCCGCCCG
TCCTCCACGCCCCCTCCACCTAACGCGCCTTGCTCTCTCCCCAACGCTGCTGGCTCTAACCTGA
GCCCGCGCTTCTGCAGAGCCTCGCTCCTCCCGCTAACCAAGCTTTCTTGCCCCCACTCTGCTGCC
TCTTGCCCGCTGCTAACGCCTCTTTGTCTATCAAGCTCACTCCCTCCCGCACAGGGCGCCCCGCCCT
25 CCTTACGGTGCACCACCAGCCCCCTGCCCCGGCTGGCCTGGCCAGGCCGCGCCCGGGATCCACAGGC
CAGCGCCACGGGGCGTCCGCGGTGTGGGTGTCCCTCGGCATGCCGCTCCCTATGCCGCGCACCTTTTCG
GGGTTAAGCCGCGATAAAGACCTTGCTTAGCTAGCAGTGCATATTGTGTAAGGTAAGGCCAGAGCCCTG
CGCGGTGGTCACTCCAGACACTGAACTGTCTTCTTTGGCAGAGAAGACGGCCACAGGGCTCGGCAGA
AGTTCCGGGGGTCTTGAGGAGGGAAGAGGGAGAGAGCCAGGGAGACCAGGGAGATGGTCGCGAGAGGAGG
30 GACTACAGGACGGGCACCAGCCATGGAGTGGCCTGAGTGTGTTGGGTTCAGCTGCCCGCAGCTGGGTTT
CTGCGGTGACGCCGAGGACTGAGGCCAGCCCTGGGGCAGCTGTGCCAGCCCCCTGTGGCGAGGCTCAC
GGGTGCCAGGGGCTGGGAGCTCAGCTCCTAGAGCAGTCTCCAGGGCTCCATCCGACAGGC
CCAGGTCACTCATGACGGGTGGCGTGGCTTGGTGTGTTTCCCCAGGCGCCAGCATGTAGCAGGCACGGT
GCCGGCTGCAGAGGTCTGGAGAGCCAGGGTGGCTGGCTGGATGCCTCTGGTCCGGGCTGTCTGGCCAGT
35 AGCCTCGGATGCCCTCATAGGGAGCCAGTGAAGCCCCAAGTGCCCTGCCTGCCTGGGGCTGATGGGTTT
CCGGCGGGTTCGAGAGGCAAGCTCTGTGCCCTGCGTGTGAGGCCCCCTTGTGTCAGACGGCTGATGGCC
ACACTGCCCTGTGGTTCATGGCCACAGCCGAGGGGCTTTTCTGCTGCTCAGATGTGACACCAAGCTGGGC
AGGTTAGTGTAGACGGTGCCACCTTTGAGACCAGAAGGAAGTTGGAAGCAGGAGTGATTTTGTATGATGT
GAGGCCTTCAGGGAATGTCTGTTCACTCTCCCCAAGAGGCTGTGGAGACTGTGGCCGCTGACCCCTG
40 AGGGCAGTTGGCCACCCGTGCCCTCCAGGGACATGGCGTGCCTGGCGGCGTAGTGCAAAGGTGCCCTCTG
CTTGCCCTGGCCTTCTGCCTCAGGGGCCCTGGGACACAGGCCCTTCTGGGCGAGTCCAAAGCCACCG
GGCAAGGAGAGGAGGACTGGGAGCCAGGAGCCAAGGCGGAGGCTGTGGCCAGGGCTGGTTCAGCCTGGCGG
GTCAATGCGGGCACCTGCGGGCTGGACAAGCTATGCTGGGTCTCTTAGGAGATGAGCTGCAGGTAGGGTG
GCTTTAGCTGGACTCGGCTGTGTCTCTTTGGAGGAAGGCTGGAGCTTGGCCGGCGTGTCCAGGTGTC
45 AGGCTCCCTGGTCTGTACGGGGCACGCCGCTGCTGTGTTGTGGTTAGAGGAGCTGGCACTATGCCTGTC
ATGGTGTGTCGGCCCCAGGTGGCGTGTGTGCTCGTGGGAGGCCTCGGTCCCCAGCCTGGGTGAGTGTGGGC
TCCCCCAGGCGGATGGTAACCTCAGGCCGCTCTTCCCGCTGTGCCCTTCTGCAGGTAACAGGATTTCT
ACCTCAACCGCGAGACCTATGCAAGGACGGTGTGGACCAACTCGCGCCCGCGGCATGGTGCCCGCGGCC
TGCCGGGCTCCACCCCTGGACGGCAGAGGCAAGGACGGACGGGAGCTCCACTGTGAATCGGCGGCACGC
50 GCCGCGAGGCTGGGACTGGTCCAGTTTGTACTGTGATAGTTTGTAGATAAAGTATTTATCATT'TTTT

SEQ ID NO:319>gi|22053032|ref|XP_046822.3| similar to FLJ00192 protein
[Homo sapiens]

MIDEIKMTDDDLFVDLGSVGVQVVLQVAAATNCKHHYGVKADIPAKYAETMDREFRKMWKWYGKKHAEY
TLERGDFLSEWRERIANSTSVIFVNNFAFGPEVDHQLKERFANMKEGGRIVSSKPFAPLNFRINSRNLSD
55 IGTIMRVVELSPLKGSVSWTGKPVSYLHTIDRTILENYFSSLNPKLREEQEAARRRQQRESKSNAATP
TKGPEGKVAGPADAPMDSGABEEKAGAATVKKPSPSKARKKKLNKKGRKMAGRKRGRPKMNTANPERKP
KKNQTALDALHAQTVSQTAASSPQDAYRSPHSPFYQLPPSVQRHSPNPLLVAPTTPPALQKLLSFQIQL
QFLAYTKTPQYKASLQELLGQEKQKNAQLLGAAQQLLSHCQAQKEEIRRLFQQLDELGVKALTYNDLIQ
AQKEISAHNQLEQSEQLEQDNRLRQSLQLLKARCEELQLDWATLSLEKLLKEKQALKSQISEQRH
60 CLELQISIVELEKSQRQELLQLKSCVPPDDALSLHLRGKALGRELEPDASRLHLELDCTKFSPLPHSS
MSPELSMNGQAAGYELCGVLSRPSSKQNTPOYLASPLDQEVVPCTPSHVGRPRLEKLSGLAAPDYTRLSP

AKIVLRRHLSQDHTVPGRPAASELHSHRAEHTKENGLPYQSPSPVPGSMKLSPOQDPRPLSPGALQLAGEKSS
 EKGLRERAYGSSGELITSLPISIPVSIPLASVVLPSRAERARSTPSPVLQPRDPSSTLEK
 QIGANAHGAGSRSLALAPAGFSYAGSVAISGALAGSPASLTPGAEPATLDESSSSGSLFATVGSRSSTPQ
 HPLLLAQPRNSLPASPAHQQLSSSPRLGGAAQGPLPEASKGDLPSDSGFSDFESEAKRRIVFTITTGAGSA
 5 KQSPSSKHSPLTASARGDCVPSHGQDSRRRRGRRKRASAGTPSLSAGVSPKRRALPSVAGLFTQPSGSPLN
 LNSMVSNINQPLEITAISSPETSLSKSSPVYQDHDQPPVLKKERPLSQTNGAHYSPLTSDEEPGSEDEPS
 SARIERKIATISLESKSPPKTLENGGGLAGRKPAPAGEPVNSSKWKSTFSPISDIGLAKSADSPLQASSA
 LSQNSLFTFRPALEEPSADAKLAAHPRKGFPGSLSGADGLSPGTNPANGCTFGGGLAADLSLHFSFDGAS
 LPHKGPEAAGLSSPLSFPSQRGKEGSDANPFLSKRQLDGLAGLKEGSGRKEAGEGGLPLCGPTDKTPLL
 10 SGKAAKARDREVDLKNHNLFIISAAVPPGSLSGPGLAPAASSAGGAASSAQTHRSFLGPFPPGPQFAL
 GPMQLQANLGSVAGSSVLQSLFSSVPAAGLVHVSSAATRLTNSHAMGSFSGVAGGTGGVFNHVPASAS
 AHPFGARVGRGAACGSATLGPSPLQAAASASASSFQAPASVETRPPPPPPPPPPPLPPPAHLGRSPAGPP
 VLHAPPPNNAALPPPTLLASNPEPALLOSLASLPPNQAFLLPTSAASLPPANASLSIKLTSPLHKGARP
 SFTVHHQPLPRLALAQAAPGIPQASATGPSAVVWSLGMPPPYAAHLSGVKPRBottom of Form

SEQ ID NO:320

Novel (maps to chromosome 4)

Novel (Chromosome 4) >

GGATAAAAAGAGTGGAAATGGGGATTTCCAGGTGCTCCCCTGGTTCATCTAGGCACCAGAGAGCTGCACTAGCAG
 20 GTCTATCATGAATCTCCTTGGAAATGCTCATTTTTAGTCCTACTTGATGTGTCTGTTCTGGAAATGCAGTATTTT
 TAATGTATCTCAACAAAATATTTTATGATTAGTAAGCTTATCTTATATAAAGGACAATTTTTTTCCTTTTTCAC
 AGGTTCTAATAATTTTTTATTTAATAATTAGATCTATTAGATTTTATTCATAACTGTGGTAGTTGAAGTACCTTC
 TAAGCTGAGTTTCAGATTTGAGAATAAACCTTGGGGTATCATTACAGAAAATTTGTCTCAATCTGCTTTGTATTT
 25 GAAAGATATGAGATTCCTGAATTATATCTTACAGACTAGTCCCCAAAAGAATACGTGTTTCTTACCTTTAAT
 TTCTCATGGTAGTTAGTCTGTGAAT

SEQ ID NO:321

Novel peroxidase-like/ melanoma antigen

>GL2-86-4M13F Direction: sense

CGACCTGGCCAGCCACCGCGGCCTGCTGCGGCAGGGCATCGTGACGCGTCCGGGAAGCCGCTGCTCCCCTTCGC
 30 CACCGGGCCGCCACGGAGTGCATGCGGGACGAGAACGAGAGCCCCATCCCCTGCTTCCCTGGCCGGGGACCACCG
 CGCCAACGAGCAGCTGGGCCTGACCAGCATGCACACGCTGTGGTTCCGCGAGCACAACCGCATTGCCACGGAGCT
 GCTCAAGCTGAACCCGCACTGGGACGGCGACACCATCTACTATGAGACCAGGAAGATCGTGGGTGCGGAGATCCA
 GCACATCACCTACCAGCACTGGCTCCCGAAGATCCTGGGGGAGGTGGGCATGAGGACGCTGGGAGAGTACCACGG
 35 CTACGACCCCGGCATCAATGCTGGCATCTTCAACGCCTTCGCCACCGCGGCCTTCAGGTTTGGCCACACGCTTGT
 CAACCCACTGCTTTACCGGCTGGACGAGAACTTCCAGCCCATTGCACAAGATCACCTCCCCCTTCACAAAGCTTT
 CTCTCTCCCCTTCGGATTG

SEQ ID NO:322

>GL2_93_2B08_G3F1 Direction: sense

CGACCTGGCCAGCCACCGCGGCCTGCTGCGGCAGGGCATCGTGACGCGTCCGGGAAGCCGCTGCTCCCCTTCGC
 40 CACCGGGCCGCCACGGAGTGCATGCGGGACGAGAACGAGAGCCCCATCCCCTGCTTCCCTGGCCGGGGACCACCG
 CGCCAACGAGCAGCTGGGCCTGACCAGCATGCACACGCTGTGGTTCCGCGAGCACAACCGCATTGCCACGGAGCT
 GCTCAAGCTGAACCCGCACTGGGACGGCGACACCATCTACTATGAGACCAGGAAGATCGTGGGTGCGGAGATCCA
 45 GCACATCACCTACCAGCACTGGCTCCCGAAGATCCTGGGGGAGGTGGGCATGAGGACGCTGGGAGAGTACCACGG
 CTACGACCCCGGCATCAATGCTGGCATCTTCAACGCCTTCGCCACCGC

SEQ ID NO:323

Novel peroxidase-like/ melanoma antigen>

CGACCTGGCCAGCCACCGCGGCCTGCTGCGGCAGGGCATCGTGACGCGTCCGGGAAGCCGCTGCTCCCCTTCGC
 50 CACCGGGCCGCCACGGAGTGCATGCGGGACGAGAACGAGAGCCCCATCCCCTGCTTCCCTGGCCGGGGACCACCG
 CGCCAACGAGCAGCTGGGCCTGACCAGCATGCACACGCTGTGGTTCCGCGAGCACAACCGCATTGCCACGGAGCT
 GCTCAAGCTGAACCCGCACTGGGACGGCGACACCATCTACTATGAGACCAGGAAGATCGTGGGTGCGGAGATCCA
 GCACATCACCTACCAGCACTGGCTCCCGAAGATCCTGGGGGAGGTGGGCATGAGGACGCTGGGAGAGTACCACGG
 55 CTACGACCCCGGCATCAATGCTGGCATCTTCAACGCCTTCGCCACCGCGGCCTTCAGGTTTGGCCACACGCTTGT
 CAACCCACTGCTTTACCGGCTGGACGAGAACTTCCAGCCATTGCACAAGATCACCTCCCCCTTCACAAAGCTTT
 CTCTCTCCCCTTCGGATTGTGAATGAGGGCGGCATCGATCCGCTTCTCAGGGGGCTGTTCCGGGGTGGCGGGGAA
 AATGCGTGTGCCCTCGCAGCTGCTGAACACGGAGCTCACGGAGCGGCTGTTCTCCATGGCACACACGGTGGCTCT
 GGACCTGGCGGCCATCAACATCCAGCGGGGCCGGGACCACGGGATCCACCCCTACCACGACTACAGGTCTACTG
 60 CAATCTATCGCGGCACACACGTTTCGAGGACCTGAAAAATGAGATTAAAAACCCTGAGATCCGG

SEQ ID NO:324

>gi|20533845|ref|XM_056455.2| Homo sapiens Melanoma associated gene
(D2S448), mRNA

5 AGCCGGCCGTGGTGGCTCCGTGCGTCCGAGCGTCCGTCCGCGCCGTGCGCCATGGCCAAAGCGCTCCAGGG
GCCCCGGGCGCCGCTGCCTGTTGGCGCTCGTGCTGTTCTGCGCCTGGGGGACGCTGGCCGTGGTGGCCCA
GAAGCCGGGCGCAGGGTGTCCGAGCCGCTGCCTGTGCTTCCGCACCACCGTGCCTGTCATGTCATCTGCTG
CTGGAGGCCGTGCCCCCGTGGCGCCGACACCTCCATCCTAGATCTTCGCTTTAACAGAATCAGAGAGA
10 TCCAACCTGGGGCATTTCAGGCGGCTGAGGAACCTGAACACATTGCTTCTCAATAATAATCAGATCAAGAG
GATACCTAGTGGAGCATTTGAAGACTTGGAAAATTTAAATATCTCTATCTGTACAAGAATGAGATCCAG
TCAATTGACAGGCAAGCATTTAAGGGACTTGCCTCTCTAGAGCAACTATACCTGCACCTTTAATCAGATAG
AAACTTTGGACCCAGATTTCGTTCCAGCATCTCCCGAAGCTCGAGAGGCTATTTTTGCATAACAACCGGAT
TACACATTTAGTTCCAGGGACATTTAATCACTTGGAACTCTATGAAGAGATTGCGACTGGACTCAAACACA
CTTCACTGCGACTGTGAAATCCTGTGGTTGGCGGATTGCTGAAAACCTACGCGGAGTCGGGGGAACGCGC
15 AGGCAGCGGCCATCTGTGAATATCCAGACGCATCCAGGGACGCTCAGTGCCAACCATCACCCCGGAAGA
GCTGAACTGTGAAAAGGCCCCCGGATCACCTCCGAGCCCCAGGACGCAGATGTGACCTCGGGGAACACCGTG
TACTTCACTGCGAGAGCCGAAGGCAACCCCAAGCCTGAGATCATCTGGCTGCGAAACAATAATGAGCTGA
GCATGAAGACAGATTCCCGCCTAACTTGTGAGCATGGGACCTGATGATCCAGAACACACAGGAGAC
AGACCAGGTATCTACCAGTGCATGGCAAAGAACGTGGCCGGAGAGGTGAAGACGCAAGAGGTGACCCTC
20 AGGTACTTCGGGTCTCCAGCTCGACCCACTTTTGTAAATCCAGCCACAGAATACAGAGGTGCTGGTTGGGG
AGAGCGTTCACGCTGGAGTGCAGCGCCACAGGCCACCCCGCGCGGATCTCCTGGACGAGAGGTGACCG
CACACCTTGGCAGTTGACCCGCGGGTGAACATCAGCCTTCTGGCGGGCTTTACATACAGAACGTCGTA
CAGGGGACAGCGGAGTAGTATGCGTGCTCTGCGAACAAACATTGACAGCGTCCATGCCACCGCTTTCA
TCATCGTCCAGGCTCTTCTCAGTTCACTGTGACGCCTCAGGACAGAGTCTGTTATTGAGGGCCAGACCGT
25 GGATTTCCAGTGTGAAGCCAAGGGCAACCCGCGCCCGTTCATCGCCTGGACCAAGGGAGGGAGCCAGCTC
TCCGTGGACCGGCGGCACCTGGTCTGTCTATCGGGAACACTTAGAATCTCTGGTGTGCCCCCAGACACC
AGGGCCAGTACGAATGCCAGGCTGTCAACATCATCGGCTCCAGAAAGTCTGGGCCACCTGACTGTGCA
GCCAGAGTCAACCCAGTGTTCGCGAGCATTCCAGCGACACAACAGTGGAGGTGGGCGCCAATGTGCAG
CTCCCGTGCAGCTCCAGGGCGAGCCGAGCCAGCCATCACCTGGAACAAGGATGGGGTTTCAGGTGACAG
30 AAAGTGGAAAAATTTACATCAGCCCTGAAGGATTCTTGACCATCAATGACGTTGGCCCTGCAGACGCAGG
TCGCTATGAGTGTGTGGCCCGGAACACCATTTGGGTGGCCCTCGGTGAGCATGGTGTCTAGTGAATGTT
CCTGACGTCACTCGAAATGGAGATCCGTTTGTAGTACCTCCATCGTGGAAAGCGATTGCGACTGTTGACA
GAGCTATAAACTCAACCCGAACACATTTGTTTGACAGCCGTCCTCGTTCTCCAAATGATTTGCTGGCCTT
GTTCCGGTATCCGAGGGATCCTTACACAGTTGAACAGGCACGGGCGGGAGAAATCTTTGAACGGACATTG
35 CAGCTCATTTCAGGAGCATGTACAGCATGGCTTGATGGTTCGACCTCAACGGAACAAGTTACCACTACAACG
ACCTGGTGTCTCCACAGTACCTGAACCTCATCGCAAACCTGTTCGGGCTGTACCGCCACCGGCGCGTGAA
CAACTGCTCGGACATGTGCTTCCACCAGAAGTACCGGACGCACGACGGCACCTGTAAACAACCTGCAGCAC
CCCATGTGGGGCGCCTCGCTGACCGCCTTCGAGCGCCTGCTGAAATCCGTGTACGAGAATGGCTTCAACA
CCCCCTCGGGGCATCAACCCCCACCGACTGTACAACGGGCACGCCCTTCCCATGCCGCGCCTGGTGTCCAC
40 CACCTGATCGGGACGGAGACCGTCACACCCGAGCAGTTCACCCACATGCTGATGAGTGGGGCCAG
TTCTTGGACACGACCTCGACTCCACGGTGGTGCCCTGAGCCAGGCACGCTTCTCCGACGGACGACACT
GCAGCAACGTGTGTGAGCAACGACCCCCCTGCTTCTGTCTGTCATGATCCCCCAATGACTCCCGGGCCAG
GAGCGGGGCCCCGTGCATGTTCTTCGTGCGCTCCAGCCCTGTGTGCGGCAGCGGCATGACTTCGCTGCTC
ATGAACTCCGTGTACCCGCGGGAGCAGATCAACCAGCTCACCTCCTACATAGACGCATCCAACGTGTACG
45 GGAGCACGGAGCATGAGGCCCGCAGCATCCGCGACCTGGCCAGCCACCGCGGCTGCTGCGGCAGGGCAT
CGTGACGCGGTCCGGGAAGCCGCTGCTCCCTTCGCCACCGGGCCGCCCCACGGAGTGCATGCGGGACGAG
AACGAGAGCCCCATCCCTGCTTCTGGCCGGGGACACCGCGCCAACGAGCAGCTGGGCCTGACCAGCA
TGCACACGCTGTGGTTCCGCGAGCACAACCGCATTGCCACGGAGCTGCTCAAGCTGAACCCGCACTGGGA
CGGCGACACCATCTACTATGAGACCAGGAAGATCGTGGGTGCGGAGATCCAGCACATCACCTACCAGCAC
50 TGGCTCCCGAAGATCCTGGGGGAGGTGGGCATGAGGACGCTGGGAGAGTACCACGGCTACGACCCCGGCA
TCAATGCTGGCATCTTCAACGCCTTCGCCACCGCGGCCCTTCAGGTTTGGCCACACGCTTGTCAACCCACT
GCTTTACCGGCTGGACGAGAACTTCCAGCCATTGCACAAGATCACCTCCCCCTTCACAAAGCTTTCTTC
TCTCCCTTCGGGATTGTGAATGAGGGCGGCATCGATCCGCTTCTCAGGGGGCTGTTCCGGGTGGCGGGGA
AAATGCGTGTGCCCTCGCAGCTGCTGAACACGGAGCTCACGGAGCGGCTGTTCTCCATGGCACACACGGT
55 GGCTCTGGACCTGGCGGCCATCAACATCCAGCGGGGCGGGACCACGGGATCCCACCTACCACGACTAC
AGGGTCTACTGCAATCTATCGGCGGCACACACGTTTCGAGGACCTGAAAAATGAGATTAAAAACCTGAGA
TCCGGGAGAACTGAAAAGGTTGTATGGCTCGACACTCAACATCGACCTGTTTCCGGCGCTCGTGGTGGGA
GGACCTGGTGCCTGGCAGCCGGCTGGGCCCCACCTGATGTGTCTTCTCAGCACACAGTTCAAGCGCCTG
CGAGATGGGGACAGGTTGTGGTATGAGAACCCTGGGGTGTCTCCCCGGCCAGCTGACTCAGATCAAGC
60 AGACGTGCTGGCCAGGATCCTATGCGACAACCGCGACAACATCACCCGGGTGCAGAGCGACGTGTTTCAG
GGTGGCGGAGTTCCTCAGCGCTACGGCAGCTGTGACGAGATCCCCAGGGTAGACCTCCGGGTGTGGCAG

5 GACTGCTGTGAAGACTGTAGGACCAGGGGGCAGTTCAATGCCTTTTCTATCATTTCCGAGGCAGACGGT
CTCTTGAGTTCAGCTACCAGGAGGACAAGCCGACCAAGAAAACAAGACCACGGAAAAATACCCAGTGTG
GAGACAGGGGGAACATCTCAGCAACAGCACCTCAGCCTTCAGCACACGCTCAGATGCATCTGGGACAAAT
10 GACTTCAGAGAGTTTGTCTTGAAATGCAGAAAGACCATCAGACACCTCAGAACACAGATAAAGAACTTG
AATCACGGCTCAGTACCACAGAGTGCCTGGATGCCGGGGGGAATCTCACGCCAACCAACCAAGTGGAA
AAAAGATGCATGCACCATTGTGTAATGCAAAGACGGGCAGGTACCTGCTTCGTGGAAGCTTGCCCCCT
GCCACCTGTGCTGTCCCGTGAACATCCCAGGGGCTGTGTCCAGTCTGCTTACAGAAGAGGGCGGAGG
AAAAGCCCTAGGCTCCTGGGAGGCTCCTCAGAGTTTGTCTGCTGTGCCATCGTGAGATCGGGTGGCCGAT
15 GGCAGGGAGCTGCGGACTGCAGACCAGGAAACACCCAGAACTCGTGACATTTTATGACAACGTCCAGCTG
GTGCTGTTACAGAAGGCAGTGCAGGAGGCTTCCAACCAGAGCATCTGCGGAGAAGGAGGCACAGCAGGTG
CCTGAAGGGAAGCAGGCAGGAGTCTTAGCTTACAGTTAGACTTCTCAGGTTTTTATTTAATTTCTTTTAA
ATGAAAAATTGGTGCTACTATTAAATTGCACAGTTGAATCATTTAGGCGCCTAAATTGATTTTGCCTCCC
AACACCATTTCTTTTAAATAAAGCAGGATACCTCTATATGTTCAGCCTTGCCCTGTTTTCAGATGCCAGGAG
20 CCGGCAGACCTGTACCCCGCAGGTGGGTGAGTCTTGGAGCTGCCAGAGGGGCTACCGAAATCGGGGTT
CCATCACAAGCTATGTTTAAAAAGAAAATTGGTGTGTTGGCAAACGGAACAGAACCTTTGATGAGAGCGTT
CACAGGGACACTGTCTGGGGGTGCAGTGCAGCCCCCGGCTCTTCCCTGGGAACCTCTGAACCTCCTCCT
TCCTCTGGGCTCTCTGTAACATTTACCACACGTCAGCATCTAATCCCAAGACAAACATTCCCCTGCTC
GAAGCAGCTGTATAGCCTGTGACTCTCCGTGTGTGAGCTCCTCCACACCTGATTAGAACATTCTAAGC
CACATTTAGAAACAGGTTTGTCTTTCAGCTGTCACTTGCACACATACTGCCTAGTTGTGAACCAATGTGA
25 AAAACCTCCTTCATCCCATTTGTGTATCTGATACCTGCCGAGGGCCAAGGGTGTGTGTTGACAACGCCGC
TCCAGCCGGCCCTGGTTGCGTCCACGTCCCTGAACAAGAGCCGCTTCCGGATGGCTCTTCCCAAGGGAGG
AGGAGCTCAAGTGTGCGGAAGTGTCTAACTTCAGGTTGTGTGAGTGCCTTAAAAAAGAGTAAAGTAA
ATCCCTATACCTCATTTGTATTTTAAATGCGTGATGTTTATGAAATTGTGTCCATTTTATAGGTATT
AGATATGGCAGAAAAACCATTTCCACTATGCAAAGTTCTTTTAGACGTCAGTGAAAAATCAACTCTCATAC
30 CTCATGGTCTCTCTTTAATTGACCAAAACCTTCCATTTTCTCTAAATACAAAGCGATCTGTGTTCTGAG
CAACCTTTCCCGAACACACAGCTTCAGTGCAGCAGCTGACCTGAGTATCCACCATGTGCCAGGCACAG
TGCTGGGCACACGAGGCACCAAGGTCCGGGCCACCTGCCCGCAGCAAGGCCAGCTGAGGTGGTGGAGGG
AGCCCTGAGGTGAGGGGCCGTTTCGGTTCAGGTTGGCAGGTGTCCAGCACTGGGGTATGGCGTCCAGGC
TTCCATGGGGTGGGGGAGGCCAGCTTCTTCTGACAGGATGGGCGCATACAGTGCCTGGTGTGATTTGTG
35 CACAACCCGTGTTCCAGGTGCACATCTCCCAAGGAGACACCCAGACCTTCCAGCACGGGCCGGCCAAG
TTGCTGCGGCGGAGGCAGCATTTTCAGCTGTGAGGAAGGTTCATGGATTTCATGTGTTTATCTGTAAAAAT
GGTTGTCTTAACCTCTTAACCTCATATTGGTAAAGTGATTGATAAAAAATTGGTTGGTGTTCATGACATGT
GGACTTCTTTTGAATAGCAAGTCAAATGTAGTGACCAAAATTGTGGAAGAGATTCTGTCAAATAGGAAA
40 TGTGTAAGTTTCGTCTAAAGCTGATGGTTATGTAAGTTGCTCAGGCACTCAGATGACAGCAGATTCTGGG
TTCTGGGAGTGTCTGTGCTCTTACATGCCCTGGAGGCCATGGTCTCAGTGTGAGGCGGCACACCT
GTAGCACACCTGCGTAATGTGCGGTCTGGGCCAGTCACAAGGAATTGTGTTGTCTAAGCCAAAGGGGAA
GCTGACTGTGATTTACCAAAAAAATTCTGTAAATTCAAACCAAAATGTCTGCGGAATCACCAGTTTGATA
CTCTCTGTAATCAGAACAGTGGGCAGTGCCTGGGTGAACGTGTCTAGCAGCCACTGTGCGGGATCGCTGT
AACAGGAGTGAATGTACATATTTATTTACTTTTCTAACTGCTCCAACAGCCAAATGCCTTTTTTATGAC
45 CATTGTATTAGTTTATTACCAAAAGAAATGTTTGACATTTGTAATGATGCCTTTCAGTTCAAATAAATGG
GTCACATTTTCAAATGG

SEQ ID NO:325

>gi|20533846|ref|XP_056455.2| similar to Peroxidasin [Homo sapiens]
45 MAKRSRGPGRCLLALVLFCAWGLAVVAQKPGAGCPSRCLCFRTTVRCMHLLLEAVPAVAPQTSILDRL
FNRIREIQPGAFRRRLNLTLLNNNQIKRIPSGAFEDLENLKYLYLYKNEIQSIDRQAFKGLASLEQLY
LHFNQIETLDPDSFQHLPLKLERLFLHNNRI THLVPGTFNHLESMKRLRLDSNTLHCDCEILWLADLLKTY
AESGNAQAAAICEYPRRIQGRSVATITPEELNCERPRITSEPQDADVTSGNTVYFTCRABGNPKPEIWL
RNNNELSMKTD SRLNLLDDGTLMIQNTQETDQGIYQCMKNVAGEVKTQEVTLRYFGSPARPTFVIQPN
50 TEVLVGESVTLCSATGPPPRISWTRGDRTPLPVDPVNITPSGGLYIQNVVQGSGEYACSATNNIDS
VHATAFIIVQALPQFTVTPQDRVVEGQTVFQCEAKGNPPVIAWTKGGSQLSVDRRLVLSSGTLRIS
GVALHDQGGYECQAVNIIGSQKVVAHLTVQPRVTFASIPSDTTVEVGANVQLPCSSQGEPEPAITWNK
DGVQVTESGKFHISPEGFLTINDVGPADAGRYECVARNTIGSASVSMVLSVNVDPVSRNGDPFVATSIVE
AIATVDRAINSTRTHLFDSDRPRSPNDLLALFRYPRDPYVEQARAGEIFERTLQLIQEHVQHGLMVDLNG
55 TSYHYNDLVSPQYLNLIANLSGCTAHRRVNCSMDCFHQKYRTHDGTENNLOHPMWGASLTAERLLKSV
YENGFNTPRGINPHRLYNHALPMPRLVSTLLIGTETVTPDEQFTHMLMQWGQFLDHDLDSTVVALSQAR
FSDGQHCSNVCSNDPPCFVSMIPPNDSDRARSARGARCMFFVRSSPVCSSGMTSLLMNSVYPREQINQLTSYI
DASNVYSGSTEHEARSIRDLASHRGLLRQGIQVRSQKPLLPFATGPPTECMRDENESPIPCFLAGDHRANE
QLGLTSMHTLWFREHNRIATELLKLNPHWDGDTIYYETRKIVGAEIQHITYQHWLPKILGEVGMRTLGEY
60 HGYPDGINAGIFNAFATAFRFGHTLVNPLLYRLDNEFQPIAQDHLPLHKAFSSPFRIVNEGGIDPLLRG
LFGVAGKMRVPSQLLNTLTERLFSMAHTVALDLAAINIQRGRDHGIPPYHDYRVYCNLSAAHTFEDLKN

EIKNPEIREKLKRLYGSTLNIDLFPALVVEDLVPGSRLGPTLMCLLSTQFKRLRDGDRLLWYENPGVFSFA
QLTQIKQTSLARILCDNADNITRVQSDVFRVAEFPHGYGSCDEIPRVDLRVWQDCCEDCRTGQFNAFSY
HFRGRRSLSEFSYQEDKPTKKTRPRKIPSVGRQGEHLNSTSAFSTRSDASGTNDFREBVLQMKTITDLR
TQIKKLESRLSTTECV DAGGESHANNTKWKKDACTICECKDGQVTCFVEACPPATCAVPVNIPGACCPVC
LQKRAEEKP

SEQ ID NO:326

Novel WD40/SOCS box protein

>GL2_81_2_M13F Direction: anti-sense

GGCTTGTAACCTTGCCACGGTGGCAGATGACAGACTCCTCAGGATCTGGGCCCTGGAAGTGAAGAACTCCCAAGTTGC
ATATTGCATCCTATNACCAATGGGCTTTGCTGCACATTTTTTCCACATGGTGGAGTCATTGCCACAGGGACAAGA
GATGGCCACNTCCAGTTCTGGACAGCTCCTAGGGTCTGTCTCACTGAAGGCACTTATGCCGGAAGCCCTTCG
AAGTTTCTTAACAACCTACCAAGTCTTAGCACTGCCTATCCCCAAGAAATGAAAGAGTTCTCTCACATACAGGAC
TTTGCTTAAAGCAACACAGCATCTTGCTGCTTTTGGTGAGCAGGTAAATCGTCTGTCAAGGGAGTTGCTGGAA
TAATGGGGCCAAACATCTGGTCTTGCAATTGAAATAGCACTTTCTTTGGGATTGTGAATAGAAATGTAGCAAAACCAG
ATTCCAGTGTACTAGTCATGGATCTTTCTCTCCCTGGCATGTGAAAGTCAGCTTAGAGGAAGAGATCCACTTGCA
CCGGCACAGAGCCGCGGAGAACCAGCGACACTGGGAGCACGGTGGCAACGGCCCTGGCGGCTCTCTGTACAGCT
CGTCANGCCGAGAGAAGGCCAAAATNCGGGAACCTGATGGAATCGCCCTTAGTGAGGGTATCCAGAACTGGCAA
CATGCCANAAACGGTACCGGGTAAACTGTTCCCCCCCATCCCCAAAAACCNACAAAGGACCGGCCNCACATAA
CAACACACAAAAGCGCACANCCATCACAGA

SEQ ID NO:327

>GL2_81_3_M13R Direction: anti-sense

GGCTTGTAACCTTGCCACGGTGGCAGATGACAGACTCCTCAGGATCTGGGCCCTGGAAGTGAAGAACTCCCAATTGCA
TTTGCTTCTATGACCAATGGGCTTTGCTGCACATTTTTTCCACATGGTGGAGTCATTGCCACAGGGACAAGAGAT
GGCCACGTCCAGTTCTGGACAGCTCCTAGGGTCTGTCTCACTGAAGCACTTATGCCGGAAGCCCTTCGAAGT
TTCTTAACAACCTACCAAGTCTTAGCACTGCCAATCCCCAAGAAATGAAAGAGTTCTCTCACATACAGGACTTTT
TTAAGCAACACCACATCTTGCTGCTTTTGTAG

SEQ ID NO:328

>GL2_81_2_M13R Direction: sense

CGGNTCTGTTGCCGTGCAAGTGGAAATCATCTTACCTCTAAGACATGACTTTTACATGCCAGGGGAGAGAAAGATC
CATGACTAGTACACTGGCAATCTGGTTTTGCTACATTCTATTACAATCCCAAGAAATGCTATTTCAATGCAAG
ACCAGATGTTTGGCCCATTTATCCAGCAACTCCCTTTGACAGGACGATTTACCCTGCTACAACAGAAGCACACAG
ATGTTGGTGTGCTTAAAAAAGTCTGTATGTGAGGAACTCTTTCACTTTTCTTGGGGATTGGCCACGTGCTATG
GACTTAGGTAAGTTGTTTCAGGAAACTTCGCAAGGGCTTTCCGGCATAAGTGCTATCAGTGAGGACAGGACCCCTAG
TGAGCTGTCCAGAACTGGACGTGGCCATCTCTTGCTCCCTGTGGCCAATGACTCCACCATGTGGAATACATGTNCA
GCAAAGCCCATTTGGTCATAAGAACAAAATGCAATCGGGAGTTNCCAGGTCCAGGGGGCCGAGATCTTGAGGGAGG
TCNTGTGCAATCTGCCAGCGTGGCAAGGGTACCAAGCCCCGANAAAACCCAGAGCAACAAGGTGGGTCTCAGA
ATNAGAATCAAGCGGGCCGGTCCGANACTAGCCTTCGCAGGGGTCTGAAATACTCACCAAAGGNCGGCAAATA
AGGGGGGGGCACATTCTGTTTCAAAACCTCGCCCCGAAAAGACNCCTAGGATGAANGGCTACAAATGACCAGCGTGA
GAAACACTGGCCNCGCAGACCCCCGGGGTGAC

SEQ ID NO:329

Novel WD40/SOCS box protein>

CGGCTCTGTTGCCGTGCAAGTGGAAATCATCTTACCTCTAAGACATGACTTTTACATGCCAGGGGAGAGAAAGATC
CATGACTAGTACACTGGCAATCTGGTTTTGCTACATTCTATTACAATCCCAAGAAATGCTATTTCAATGCAAG
ACCAGATGTTTGGCCCATTTATCCAGCAACTCCCTTTGACAGGACGATTTACCCTGCTACCAAGAAGCACACAAG
TGTGGTGTGCTTAAAAAAGTCTGTATGTGAGGAACTCTTTCACTTTTCTTGGGGATTGGCAGTGCTAGGACTTG
GTAAGTTGTTAGGAAACTTCGAAGGGCTTTCCGGCATAAGTGCTTCAGTGAGGACAGGACCCCTAGGAGCTGTCCA
GAACTGGACGTGGCCATCTCTTGTCCTGTGGCAATGACTCCACCATGTGGAATAATGTGCAGCAAAGCCCATTT
GGTCATAGGAGCAATGCAATGGGAGTTTTCAGTTCCAGGGCCAGATCCTGAGGAGTCTGTCACTGCCACCGT
GGCAAGGTACAAGCCCCCG

SEQ ID NO:330

>gi|20149658|ref|NM_018639.2| Homo sapiens CS box-containing WD protein
(LOC55884), mRNA

CGGACGCGTGGGGCCCGGGGGGCTGAGGCGCCCGCCGCTGCCGCGGGGGCCGCTCGCGTCTCTCCATGGA
GGCCGGAGAGGAACCGCTGCTGCTGGCCGAACCTCAAGCCCGGGCGCCCCACCAGTTTGATTGGAAGTCC
AGCTGTGAAACCTGGAGCGTCGCTTCTCCCAGATGGCTCCTGGTTTGGTCTCAAGGACACTGCA

TCGTCAAACATGATCCCCTGGCCGTTGGAGGAGCAGTTCATCCCTAAAGGGTTTGAAGCCAAAAGCCGAAG
TAGCAAAAATGAGACGAAAGGGCGGGGCAGCCCAAAAGAGAAGACGCTGGACTGTGGTCAGATTGTCTGG
GGGCTGGCCCTTCAGCCCGTGGCCCTTCCCCACCCAGCAGGAAGCTCTGGGCACGCCACCACCCCAAGTGC
CCGATGTCTCTTGCCTGGTTCTTCCGCGCCACCAAGATGTCGTGAGAGATCTGAGCTTCACACCCAGTGGC
5 AGGGCTCCTGCTTTTGAATCTTTCCGCGCCACCAAGATGTCGTGAGAGATCTGAGCTTCACACCCAGTGGC
AGTTTGATTGTTTGGTCTCCGCGTCACGGGATAAGACTCTTCGCATCTGGGACCTGAATAAACACGGTAAAC
AGATTCAAGTGTATCGGGCCACCTGCAGTGGGTTTACTGCTGTTCCATCTCCCCAGACTGCAGCATGCT
GTGCTCTGCAGCTGGAGAGAAGTCGGTCTTCTATGGAGCATGAGGTCTACACGTTAATTCGGAAGCTA
GAGGGCCATCAAAGCAGTGTGTCTCTTGTGACTTCTCCCCGACTCTGCCCTGCTTGTACGGCTTCTT
10 ACGATACCAATGTGATTATGTGGGACCCCTACACCGCGCAAAGGCTGAGGTCACTCCACCACACCCAGGT
TGACCCCGCCATGGATGACAGTGACGTCCACATTAGCTCACAGATCTGTGTGCTTCTCTCCAGAAGGC
TTGTACCTTGCCACGGTGGCAGATGACAGACTCCTCAGGATCTGGGCCCTGGAACCTGAAAACCTCCCATG
CATTGTGCTCCTATGACCAATGGGCTTGTGTCACATTTTTTCCACATGGTGGAGTCATTGCCACAGGGAC
AAGAGATGGCCACGTCCAGTCTTGACAGCTCCTAGGTCCTGTCTCACTGAAGCACTTATGCCGGA
15 GCCCTTCGAAGTTTCTTAACAACCTTACCAAGTCCTAGCACTGCCAATCCCCAAGAAAATGAAAGAGTTCC
TCACATACAGGACTTTTTAAGCAACACCACATCTTGTGCTTCTTTGTAGCAGGGTAAATCGTCTGTCAA
AGGGAGTTGCTGGAATAATGGGCCAAACATCTGGTCTTGCATTGAAATAGCATTTCTTTGGGATTGTGAA
TAGAATGTAGCAAAACCAGATTCCAGTGTACTAGTCATGGATCTTCTCTCCCTGGCATGTGAAAGTCAG
TCTTAGAGGAAGAGATTCCACTTGCACGGCAACAGAGCCTTACGTTAAATTTTTCAGTCCAGTTATGAACA
20 GCAAGTGTGAACTCTTCTGCTTGTGTTTGTGATTCAAAGTGCAGTTACTGATGTTGTTTTGATTATGCAAC
TAAGTAGGCCTCCAGAGCCTCTCTAGTGGCAGAGCAGCTCACACTCCCTCCGCTGGGAACGATGGCTTCT
GCCTAGTACCTATCCTTGTGCTTTCTGATGCAGTGGTAGCATTTGGTCAAGTCTCTCTCTGCTGTGGTCAG
AGTTGCTTCGATGTTGGCCAAAGTCTTTCTTCTGGGCTCCCTTCTGACCTGCAGGACAGTTTTCCTGG
AGCCATTTGGTATGAGGTATTAATTTAGCTTAACTAAATTACAGGGGACTCAGAGGCCGTGCTCCTGACC
25 GATCCAGACACTATTACTGGCTTTTTTTTTTTTTTTTAAACAATGGTGTGCATGTGCAGGAAATGACAAATTT
GTATGTCAGATTATACAAGGATGTATTCTTAAACCGCATGACTATTAGATGGCTACTGAGTTATCAGTG
GCCATTTATTAGCATCATATTTATTTGTATTTTCTCAACAGATGTTAAGGTACAACCTGTGTTTTCTCGA
TTATCTAAAACCATAGTACTTAAATTGAACAGTTGCAAAGATGCTTAATTGTGTAAAGAATTGGTGTA
GTCATGACTTTAGCTGATACTCTTATGTACGAGATCTGTCTCTGCTGTTTAACTTCATTGGATTAATCAG
30 CTGGTTTCAACTCTACTGCGAAACAAAAATAGCTCCTTAAAGTACTGTTCTCTCCTTCAGTGGCATGTAGT
TATCTAATCAAGACACCTCATTCAAACAAACCTGCCTTAGGAAAATTTAATATATTTTAAATTATTTTA
AAAGAAATACAACATCTTATTCTTTAGCTTTCTTAATCGGTGCTTTATGGAGGCCAGTGTAACGTTACAT
GACTCGTTGAGAAAGTTGAGGAATTTCTCTACCACTTTGTTGCTTGAAGAAAAACATGCTTTTCAAA
ATGAGAGGCTTTTCATTGAAGAAAAGAAAAAAACAACAGTTAAAAA
35 AAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO:331

>gi|8923881|ref|NP_061109.1| CS box-containing WD protein [Homo sapiens]
MEAGEEPLLLAELKPRPHQFDWKSSCETWSVAFSPDGSWFAWSQGHCIKLI PWPLEEQFIPKGFPAKS
40 RSSKNETKGRGSPKEKTLDCGQIVWGLAFSPWPSPSRKLWARHHPQVPDVSLVLATGLNDGQIKIWEV
QTGLLLNLSGHQDVVRDLSTFPGSLILVSASRDKTLRIWDLNKHGKQIQVLSGHLQWVYCCSISPDCS
MLCSAAGEKSVFLWSMRSYTLIRKLEGHQSSVVSDFSPDSALLVTASYDNTVIMWDPYTGERLRLSHHT
QVDPAMDDSDVHISSLRSVCFSPGELYLATVADRLLRILWALELKTPIAFAPMTNGLCCTFFPHGGVIAT
45 GTRDGHVQFWTAPRVLSSLKHLCKALRSFLTTYQVLALPIPKMKKEFLTYRTF

SEQ ID NO:332

Numb-like

>GL3_16A_3_M13R Direction: anti-sense
CGTCTGCCTGCCACTGGTGCAGGGCGCGACGCCTCGGGCACGTAGGCTGGCTTCCTCCGCCGCAGGGCTCTTGNCC
50 GGTAAAACCTTGTTTCATGGTGCAGCGCACCCGTCTGGCTCCGTCTGCAGGGTTTCTTGGGGGGCCCCGGGGCCCC
CACAAGGGGGACTGGGGGGCAGGGTGCCGGTTCAAGGGCCCCCTTCCTTTGTTACAGGCTTCGGTCCAATGCNG
AGAAGGAAAGGGGCGAAATTCNCTGCAGAATA

SEQ ID NO:333

>GL3_16A_4_M13R Direction: anti-sense
CGTCTGCCTGCCACTGGTGCAGGGCGCGACGCCTCGGGCACGTAGGCTGGCTTCCTCCGCCGCAGGCTCTGCCGTA
ACTTGTTTCATGGTTGCCCGCCCCGTCTGGCTCCGTCTGCAGGTTTCTGGAGGCCACCGTTGCCATCGGGCCAT
GGGCGGGCCTTCCTTCCCTTTGTTACCAGGCCTTCCGGTTCCCCAATAGCCCGGGAAGGAAGAAAACGGGGGCGG
55 AAATTTCTTGGCCAGGAGTTATTCCCNATTTCAAACAACTTGGGCCGGGGGACCCGGAGTTTCCGGAAAGCNC
ATATTTGNCAAATTCATAAGGAGAAGGGGGGGGGCCCCCAAAAATTTTCGGGCAGCCCCCTTTAATTTANGGT

TTGAGAAAGGATTTCCGGGTAATTTTACACCAAAATTTTCAACTTTGTGGGGCGCCAGGATACAGAGGATAT
TAAAAACACAAAAACGGCGGGGTANCAGGGGTNTGGGNCCTTTGGGGGGGAGAGAAGAAAAACCCGCCCCCT
TTTTGTGGGCCGNGGNTTTTATAACGCCCCCCCCGANAAAAATCCTTTTATAAAATTTCCATGCGAGCCGCGCTT
TTTTTTNGGNGTGAGAGAGGGACAGGAGCACAAATATTTTCCCCCGCCCTNTTTATTTTTTTACCGAGGGCG
5 CCCCANAGGCGTTATTGGGGACGCGGCCGTTTACAAAATAATAATGACGCTGCGAGAAAAATAGAGAAAAG
GGTGGCCCCCCCCCTGCGGAGAANAACCCGCGNGGNAGATATTACTTGGCAGGGAGCGCCCCCTTTTAT

SEQ ID NO:334

>GL3_39L_1_M13R Direction: anti-sense

CGTCTGCCTGCCACTGGTGCGGGCGCGACGCCTCGGGCACGTAGGCTGGCTTCCTCCGCCGAGGCTCTGCCGTA
ACTTGTTCATGGTGCCCGCCCCCGTCTGGCTCCGTCCGTGCAGGTTTTCTGGGGGCCCCGGGGCCCCCAACAAGGG
GGGGCCTTGGGGGGGCCAGGGTTGCCGAGAAAGCCCAAGGCAACACTTGGNGAAAGCCCAACCGGNTAGGCACAA
CCGGGGCCACTTATNGGCGANGGAGCCATNACACATTCCTTTTGGTTACCCAAGGGCANATCCGGATCCCCAAT
TTGCACCCGGANATGGAAAGCAAAAGGGGGCGCCGGAAAAAATTTCTTTGCCCGAGGATTAATTTCCCCAAATT
15 CCAAACACACCCTGTGGTGGCACGGNGACCACGGAATTCGAAAAGGCCNATTTTGGCCCATATTCTTTTAGGG
AAAGGGGGGGCCCCCCCCAAAAGAATTTTGTTCNGGCCCCCCCATATATAAA

SEQ ID NO:335

>GL3_16A_2_M13R Direction: anti-sense

CGTCTGCCTGCCACTGGTGCGTGGCGCGACGCCTCGGGCACGTAGGCTGGCTTCCTCCGCCGAGGCTCTGCCGT
TAAACTTGTTTCATTGGTGCCCGCCCCGTTCTTGGCCTCCGTCCCTTGACAGTTTTCTGGGGGCCCCGGNGGGGGCCC
CCAAACAAGGGGTGGCTTTGGAAGCCCCAACCGGTTGCCAACNCGGACACTATGGGGGGACCTTTCACATTCC
TTTTGTTACAGGCTTCGGTTCACAATTGGCCCCAAGGAAAGAAAAGGGGGCGCGGAAATTTCTTGTGCCAAGG
ATTTTATTCCCCAATTTTCAAGCGACCGTTGGGCCCGGGGACCCCGGACTCCGAAGCCATTGGCAATTCATNAG
25 GGAAAGGGNGCCCCCAAAATTTTCGCGGACCTATAAATAAGATTGAGCAGACTCGGGTAATTTATAACAAATTT
TACCACAACGCGTTGTTGGGGACCCCCGGGTATACTGCAGGTTATTTATTATAACCANAAAAGGCGGTGATATTC
CANGTAGTTGTGGAACCTGTTGTGGNGGGGAAACAAAAAACAACACCCCCCCCCCTTTTGGGGTANGGC
CGGGTGATTTTATAACAACCCCCCAGCCAAAAAGAAGCCGCTTTTNTAANAAG

SEQ ID NO:336

>GL3_39L_3_M13F Direction: N/A

CGAGGAGGCTCGCGCGCGTGCATCCCGCACCATCCCCAGCCCCAGAAACCTGCAGGACGGAGCCAGACGGGGT
GGGCACCATGAACAAGTTACGGCAGAGCCTGCGGCGGAGGAAGCCAGCCTACGTGCCCCAGGGCCGTGCGCCCCG
AACCAGTGGCAGGCAGACG

SEQ ID NO:337

>GL3_39L_3_M13R Direction: anti-sense

CGTCTGCCTGCCACTGGTGCGGGCGCGACGCCTCGGGCACGTAGGCTGGCTTCCTCCGCCGAGGCTCTGCCGGT
AACTTGTTTCATGGTGCCACCCCGTCTGGCTCCGTCCGTGCAGGTTTTCTGGGGGCTGGGGGATGGTGCGGGATGCA
40 ACGCGCGCGAGCTTCCTCG

SEQ ID NO:338

>GL3_16A_1_M13F Direction: anti-sense

CGTCTGCCTGCCACTGGTGCGGGCGCGACGCCTCGGGCACGTAGGCTGGCTTCCTCCGCCGAGGCTCTGCCGTA
45 ACTTGTTCATGGTGCCCGCCCCGTCTGGCTCCGTCCGTGCAGGTTTTCTGGGGGATGGTGCGGGATGCACGCGCGCG
AGCCTCCTCG

SEQ ID NO:339

>GL3_16A_1_M13R Direction: sense

CGAGGAGGCTCGCGCGCGTGCATCCCGNACCATCCCCAGAAACCTGCAGGACGGAGCCAGACGGGGCGGGCACC
50 ATGAACAAGTTACGGCAGAGCCTGCGGCGGAGGAAGCCAGCCTACGTGCCCCAGGCGTTCGCGCCCCGACCAGTGGC
AGGCAGACG

SEQ ID NO:340

>gi|10863898|ref|NM_004756.1| Homo sapiens numb homolog (Drosophila)-like (NUMBL), mRNA

ATGTCCCGCAGCGCGGCGGCCAGCGGCGGACCCCGAGGCCTGAGCGGCACCTGCCCCAGCCCCCTGTG
GGGCCCCGGGGCCCCCAGAAACCTGCAGGACGGAGCCAGACGGGGCGGGCACCATGAACAAGTTACGGCA
GAGCCTGCGGCGGAGGAAGCCAGCCTACGTGCCCCAGGCGTTCGCGCCCCGACCAAGTGGCAGGACGAG
60 GACGCGGTGCGGAAGGGCACGTGCAGCTTCCCGGTGAGGTACCTGGGTACGTGGAGGTAGAGGAGTCCC

GGGGAATGCACGTGTGTGAAGATGCGGTGAAGAAGCTGAAGGCGATGGGCCGAAAGTCCGTGAAGTCTGT
CCTGTGGGTGTCAGCCGATGGGCTCCGAGTGGTGGACGACAAAACCAAGGATCTTCTGGTCGACCAGACC
ATCGAAAAGGTCTCCTTTTGTGCTCCTGACCGCAACCTGGACAAGGCTTTCTCCTATATCTGTCTGTGACG
5 GGACTACCCGCCGCTGGATCTGCCACTGTTTTCTGGCACTGAAGGACTCCGGCGAGAGGCTGAGCCACGC
TGTGGGCTGTGCTTTTGCCGCCCTGCCCTGGAGCGAAAACAGCGACGGGAGAAGGAATGTGGGGTACAGGCC
GCCTTCGATGCCAGCCGACACGACTTCGCCCCGCGAGGGCTCCTTCGCCCTGTCTGGGGTGGGGCGGCTGT
CTGAGCGGAGGCCCCGGACAGAAGAAGCAGAGGACGACGCTGCCCCCTACTGTGGCTCTCTGGCCCTGC
CCAGCCTGGAGCAGCTGTCCCCGACACCGACCACCACTCCCCCTGGTGAGAAGGGTGAGGCAGGCACCCCT
GTGGCTGCAGGCACCACTGCGGGCGCCATCCCCCGGCGCCATGCACCCCCTGGAGCAGCTGGTTTCGCCAGG
10 GCTCCTTCCGTGGGTTCACAGCACTCAGCCAGAAGAACTCGCCTTTCAAACGGCAGCTGAGCCTACGGCT
GAATGAGCTGCCATCCACGCTGCAGCGCCGCACTGACTTCCAGGTGAAGGGCACAGTGCCTGAGATGGAG
CCTCCTGGTGCCGCGACAGTGACAGCATCAACGCTCTGTGACACACAGATCAGTTTATCTTTTGCCAGTG
CTGGAGCGCCAGCACCAGGGCCACCACCTGCCACAACAGGGACTTCTGCCCTGGGGTGAGCCCCTCCGTGCC
15 CCTCGAGCTGCCTTCCAGCCTGGGACACAAGCGGACACCTTCAGAGGCTGAGCGATGGCTGGAGGAGGTG
TCACAGGTGGCCAAAGGCCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC
AGCAGCAAGCAGCCTCAGTGGCCCCAGTGCCCAACCATGCTCTCTGCCCTGACAGCCTTTCCCCGCCGCCGT
GGGGCCCTTTGACGCTGCACCTGCCCAAGTGCCCGTGTTCCTGCCACCCCCACACATGCAGCCCCCTTTT
GTGCCCCCTACCCGGGCTTGGGCTACCCACCGATGCCCCGGGTGCCCGTGGTGGGCATCACACCCCTCAC
20 AGATGGTGGCAAACGCCCTTCTGCTCAGCCGCCAGCTCCAGCCTCAGCCTGCCACTCTGCTTGGGAAAGC
TGGGGCCTTCCCGCCCCCTGCCATACCCAGTGCCCCCTGGGAGCCAGGCCCGCCCTCGCCCCAATGGGGCC
CCCTGGCCCCCTGAGCCAGCGCTGCCCCAGCTCCAGAGTTGGACCCCTTTTGAGGCCAGTGGGCGGCAT
TAGAAGGCAAAGCCACTGTAGAGAAACCTCCAACCCCTTTTCTGGTGACCTGCAAAAGACATTTCGAGAT
TGAAGTGTAG

25 SEQ ID NO:341

>gi|10863899|ref|NP_004747.1| numb homolog (Drosophila)-like; numb
(Drosophila) homolog-like [Homo sapiens]
MSRSAAASGGPRRPERHLPPAPCGAPGPPETCRTEPDGAGTMNKLQRSLRRRKPAYVPEASRPHQWQADE
DAVRKGTCSFPVRYLGHVEVEESRGMHVCEDAVKKLLKAMGRKSVKSVLWVSADGLRVDDDKTKDLLVDQT
IEKVSFCAPDRNLDKAFSYICRDGTTTTRRWICHCFALAKDSGERLSHAVGCAFAACLERKQORREKECGVTA
AFDASRTSFAREGSFRLSGGGRPAEREAPDKKKAERAAAAAPTVPAGPAQPGHVSPTPATTSPEGEKGEAGTP
VAAGTTAAAIPIRRHAPLEQLVRQGSFRGFPALSQKNSPFKRQLSLRLNELPSTLQRRTDFQVKGTVPEME
PPGAGDSDSINALCTQISSFSASAGAPAGPPPTATGTSAWGEPSVPPAAAFQPGHKRTPSEAERWLEEV
SQVAKQDDQQQQQQQQQQQQQQQAAASVAPVPTMPFALQFPFAPVGFDAAPAQVAVFLPPPHMQPPF
VPAYPGLGYPPMPRVVVGITPSQMVANAFCSAAQLQCPQATLLKAGAFPPPAIPSAQSGARPRPNGA
PWPPEPAPAPAPELDPFEAQWAALEGKATVEKPSNPFSGDLQKTFEIEL

SEQ ID NO:342

40 P130Cas (Crk-associated substrate)
>GL1_43_2_1_M13F Direction: sense
CGCGGGATGTCGTACTCGTCTGCTCCGGCTGGGCGGCCTCGTATACATAGCCCTGCCCCACGCGGGTGGGCACC
ACCACCTTTGCCGGGGGCTTCGTGCCCTCCACAGTCTCGTGTGTGCCATGGACGGGGGGACCTGGTAGATGTCATGC
CCATCCCCGCAAGGTGGCACCCTGGTAAATATCTTGGGCAGGGCCTCCAGGCCCTGGGGGCACCTGGTACAGG
TCTGTGGCCGGGCTGGGAAACGGGTGATGGGGTGTCTGCTTCGAGAAGGTGGATGTCTGCTTGGC

SEQ ID NO:343

50 >GL1_43_2_1_M13R Direction: anti-sense
GCCAAGCAGACATCCACCTTCTCGAAGCAGACACCCCATCACCCGTTTCCAGCCCCGCCACAGACCTGTACCAG
GTGCCCCCAGGGCCTGGAGGCGCCCTGCCAGGATATTTACCAGGTGCCACCTTCTGCCGGGATGGGGCATGACATC
TACCAGGTCCCCGATCCATGGACACACGACATCTGGGAGGGCAGCAAGCCCCGGCAAAGGTGGGTGGTGGCCAC
CCGCGTGGGGCAGGGCTATGTATACGAGGCGCGCCAGCGGAGCAGGACGAGTACGACATCCCCGCG

SEQ ID NO:344

55 >GL1_43_3_6_M13F Direction: sense
CGCGGGATGTCGTACTCGTCCGTGCCGGTGGGCGGCCCTCGNTATACATAGCCCTGCCCCACGCGGGTGGGCAC
CACCACCTTTGCGGGGGGCTTCGTGCCCTCCCAGCTGCGTGTGTCCATGGACGGGGGGACCTGGTAGATGTCATG
CCCCATCCCGGAGTGGTGACCTGGTAAATCTCGGCAGGGCCTCCAGGCCCTGGGGGACCTGGTACAG
GTCTGTGGCCGGGGTGGTAAACGGGTGATGGGGTGTCTGCTTCGAGAAAGGTGGATGCTGTCTTGGC

SEQ ID NO:345

>GL1_43_3_6_M13R Direction: anti-sense

GCCAAGCAGACATCCACCTTCTCGAAGCAGACACCCCATCACCCGTTTCCCAGCCCCGGCCACAGACCTGTACCAG
GTGCCCCCAGGGGCGCTGGAGGCCCTGCCAGGATATTTACCAGGTGCCACCTTCTGCCGGGATGGGGCATGACAT
CTACCAGGTCCCCCGTCCATGGACACACGCAGCTGGGAGGGCACGAAGCCCCCGGCAAAGGTGGTGGTGCCAC
CCGCGTGGGGCAGGGCTATGTATACGAGGCCGCCAGCCGGAGCAGGACGAGTACGACATCCCCGC

SEQ ID NO:346

>GL1_43_4_4_M13F Direction: sense

CGCGGGATGTCTACTCGTCCCTGCTCCGGCTGGGCGGCCTCGTATACATAGCCCTGCCCCACGCGGGTGGGCACC
ACCACCTTTGCCGGGGGCTTCGTGCCCTCCCAGCTGCGTGTGTCCATGGACGGGGGGACCTGGTAGATGTCATGC
CCCATCCCCGGCAGAAGGTGGCACCTGGTAAATATCCTGGGCAGGGCCTCCAGGCCCTGGGGGCACCTGGTACAGG
TCTGTGGCCGGGCTGGGAAACGGGTGATGGGGTGTCTGCTTCGAGAAGGTGGATGTCTGCTTGGC

SEQ ID NO:347

>GL1_43_4_4_M13R Direction: anti-sense

GCCAAGCAGACATCCACCTTCTCGAAGCAGACACCCCATCACCCGTTTCCCAGCCCCGGCCACAGACCTGTACCAG
GTGCCCCCAGGGGACCTGGAGGCCCTGCCAGGATATTTACCAGGTGCCACCTTCTGCCGGGATGGGGCATGACAT
CTACCAGGTCCCCCGATCCATGGACACACGCAGCTGGGAGGGCACGAAGCCCCCGGCAAAGGTGGTGGTGCCCA
CCGCGTGGGGCAGGGCTATGTATACGAGGCCGCCAGCCGGAGCAGGACGAGTACGACATCCCCGC

SEQ ID NO:348

>GL1_43_4_6_M13F Direction: sense

CGCGGGATGTCTACTCGTCCCTGCTCCGGCTGGGCGGCCTCGTATACATAGCCCTGCCCCACGCGGGTGGGCACC
ACCACCTTTGCCGGGGGCTTCGTGCCCTCCCAGCTGCGTGTGTCCATGGACGGGGGGACCTGGTAGATGTCATGC
CCCATCCCCGGCAGAAGGTGGCACCTGGTAAATATCCTGGGCAGGGCCTCCAGGCCCTGGGGGCACCTGGTACAGG
TCTGTGGCCGGGCTGGGAAACGGGTGATGGGGTGTCTGCTTCGAGAAGGTGGATGTCTGCTTGGC

SEQ ID NO:349

>GL1_43_4_6_M13R Direction: anti-sense

GCCAAGCAGACATCCACCTTCTCGAAGCAGACACCCCATCACCCGTTTCCCAGCCCCGGCCACAGACCTGTACCAG
GTGCCCCCAGGGNCTTGGAGGCCCTGCCAGGATATTTACCAGGTGCCACCTTCTGCCGGGATGGGGCATGACAT
CTACCAGGTCCCCCGATCCATGGACACACGCAGCTGGGAGGGCACGAAGCCCCCGGCAAAGGTGGTGGTGCCCA
CCGCGTGGGGCAGGGCTATGTATACGAGGCCGCCAGCCGGAGCAGGACGAGTACGACATCCCCGC

SEQ ID NO:350

>GL1_43_LOW_3_G3F1 Direction: sense

GCGGGATGTCTACTCGTCCCTGCTCCGGCTGGGCGGCCTCGTATACATAGCCCTGCCCCACGCGGGTGGGCACCA
CCACCTTTGCCGGGGGCTTCGTGCCCTCCCAGCTGCGTGTGTCCATGGACGGGGGGACCTGGTAGATGTCATGCC
CCATCCCCGCGAGAAGGTGGCACCTGGTAAATATCCTGGGCAGGGCCTCCAGGCCCTGGGGGCACCTGGTACAGGT
CTGTGGCCGGGCTGGGAAACGGGTGATGGGGTGTCTGCTTCGAGAAGGTGGATGTCTGCTTGGC

SEQ ID NO:351

>gi|7656923|ref|NM_014567.1| Homo sapiens breast cancer anti-estrogen resistance 1 (BCAR1), mRNA

GAGGCGGCAGCTGCGCGGCGGCACCGGGCGGCTGCGGCGCGCTCGGAGCCCCGAGGCACGCGGCCCGGG
CAGCTCGGTGTGCGCCCCGCGAGAGCCGGGCCCCAGGCCCGCCGGACACCATGAACCACCTGAACGTGC
TGGCCAAAGCGCTCTATGACAATGTGGCCGAGTCCCCGGATGAGCTCTCCTTCCGCAAGGGTGACATCAT
GACGGTGTCTGGAGCAGGACACGCGAGGGCCTGGACGGCTGGTGGCTCTGCTCGCTGCATGGGCGCCAGGGC
ATCGTGCCTGGGAACCGCTCAAGATCTTGGTGGGCATGTATGATAAGAAGCCAGCAGGGCCTGGCTCCG
GCCCTCCCGCCACCCCGGCCAGCCTCAGCCTGGCCTCCATGCCCCAGCGCCTCCGGCCTCCAGTACAC
GCCCATGCTCCCCAACACCTACCAGCCCCAGCCAGACAGCGTCTACCTGGTGGCCACTCCCAGCAAGGCT
CAGCAAGGCCTCTACCAAGTCCCGGTCCCAGCCCTCAGTTCCAGTCTCCCCAGCCAAGCAGACATCCA
CCTTCTCGAAGCAGACACCCCATCACCCGTTTCCCAGCCCCGGCCACAGACCTGTACCAGGTGCCCCCAGG
GCCTGGAGGCCCTGCCAGGATATTTACCAGGTGCCACCTTCTGCCGGGATGGGGCATGACATCTACCAG
GTCCCCCGTCCATGGACACACGCAGCTGGGAGGGCACGAAGCCCCCGGCAAAGGTGGTGGTGCCACCC
GCGTGGGGCAGGGCTATGTATACGAGGCCGCCAGCCGGAGCAGGACGAGTACGACATCCCCGCGACACCT
GCTGGCCCCGGGGCCACAGGACATCTATGATGTGCCCCGGTTCCGGGGCTGCTTCCCAGCCAGTATGGC
CAGGAGGTGTATGACACACCCCATGGCTGTCAAGGTCCCAATGGCCGAGACCCGTTGCTGGAGGTGT
ATGACGTGCCCCCAGTGTGGAGAAGGGCCTGCCACCGTCCAACCACCACGCAGTCTACGACGTTCTCTCC

ATCGGTGAGCAAGGATGTGCCCGATGGCCCACTGCTGCGTGAGGAGACCTACGATGTGCCCCCGCCTTC
GCCAAGGCCAAGCCCTTTGACCCGGCCCGCACCCCACTGGTACTGGCTGCGCCCCCTCCAGACTCCCCGC
CGGCCGAGGACGTGTATGACGTGCCGCCCCCGGCTCCTGACCTCTACGACGTGCCCCCTGGCTTGCGGCG
5 GTGGTTCGACAGTGGTGTGTATGCGGTGCCCTCCCCAGCTGAACGTGAAGCCCCGGCAGAGGGCAAGCGCC
TGTCGGCCTCCAGCACCGGCAGCACACGCAGCAGCCAGTCTGCGTCCTCCTTGGAGGTGGCAGGGCCGCGG
CCGGGAACCCCTGGAGCTGGAAGTTGCTGTGGAGGCCCTGGCAGGGCTGCAGCAGGGTGTGAGCGCCACC
GTTGCCACCTTCTGGACCTGGCAGGCAGCGCCGGTGCAGCTGGGAGCTGGCGTAGCCCCCTCTGAGCCAC
10 AGGAGCCGCTGGTGCAGGACCTGCAGGCTGCTGTGGCCGCTGTCCAGAGTGCCGTCCACGAGCTGTTGGA
GTTTGGCCCGCAGCGCGGTGGGCAATGCTGCCACACATCTGACCGTGCCCTGCATGCCAAGCTTAGCCGG
CAGCTGCAGAAGATGGAGGACGTGCACCAGACGCTGGTGGCACATGGTCAGGCCCTCGACGCTGGCCGGG
GAGGCTCTGGAGCCACCCTTGAGGACCTGGACCGGCTGGTGGCCTGCTCGCGGGCTGTGCCCGAGGACGC
CAAGCAGCTGGCCTCCTTCTGACGGCAATGCCCTACTGCTCTTACAGACGGACCAAGGCCACTGCCCCG
GGCCCTGAGGGGGGTGGCACCCCTGCACCCCAACCCCACTGACAAGACCAGCAGCATCCAGTCACGACCC
15 TGCCCTCACCCCTAAGTTACCTCCCAGGACTCACCAGATGGGCGATACGAGAACAGCGAGGGGGGCTG
GATGGAGGACTATGACTACGTCCACCTACAGGGGAAGGAGGAGTTTGAGAAGACCCAGAAGGAGCTGCTG
GAAAAGGGCAGCATCACGCGGCAGGGCAAGAGCCAGCTGGAGTTGCAGCAGCTGAAGCAGTTTGAACGAC
TGGAACAGGAGGTGTACGGCCCATAGACCAGACCTGGCCAACTGGACGCCAGCCCAACCCCTGGCCCC
GGGGCGAACAGGCGGCCTGGGGCCCTCGGACCGGCAGCTGCTGCTCTTCTACCTGGAGCAGTGTGAGGCC
20 AACCTGACCACACTGACCAACGCCGTGGACGCCTTCTTTACCGCCGTGGCCACCAACCAGCCGCCCAAGA
TCTTTGTGGCGCACAGCAAGTTCGTATCCTCAGCGCCCAAGCTGGTGTTCATCGGGGACACACTGTC
ACGGCAGGCCAAGGCTGCTGACGTGCGCAGCCAGGTGACCCACTACAGCAACCTGCTGTGCGACCTCCTG
CGCGCATCGTGGCCACCACCAAGCCGCTGCCTTGACATGCCATCGCCTTCCGCGGCCAGGACATGG
TGGAGAGGGTCAAGGAGCTGGGCCACAGCAGCTTCCGCCGCGTCTTAGGCCAGCTGGCAGCCGCG
25 CTGAGGGTGGTGACCCAGGAGGGAGGCAGGGGAGGGGTGCGCGGTCCAGCTCCCTGGCTCCCATGTC
AAGAGTCGCTGTGCCACAGGCTTAGGGACAGGACCCAGCTCTGCGTCGGTCTGGTGGCCTGGATGCCC
AGGAATCTGTATATATTTATGGCCGGGCAGGGTGTGGGGCCATGCCCTCCTCAGGAGCCGAAGCCAGGGG
CCGCACTGGCCTTCCCCAGCATGCACCACGGGCCCGGTTGGGTACCAGACGGGGCTGGAGTGTGAGGG
TCCTGCAGCCTGCAGGACCTCGTGCCACCCCGAGGGCTGAGCCTGGTCCCACGAGGGTGGCGTGTCCCT
30 GACAGGGCCAGTGCAGTTTGGTGTGTCTCCGCTTACCAGGAGAAGAACCTGAAGAACTATTTTCGTT
ATTGGTTTTCCAATCATTGACTAAGAGTCTCCATTTAAATAAAGTTTTTAAAGGAA

SEQ ID NO:352

>gi|7656924|ref|NP_055382.1| breast cancer anti-estrogen resistance 1; Crk-
35 associated substrate p130Cas [Homo sapiens]
MNHLNLAKALYDNVAESPDELSFRKGDIMTVLEQDTQGLDWWLCSLHGRQGIIVPGNRLKILVGMYDKK
PAGPGSGPPATPAQPQPLHAPAPPASQYTFMLPNTYQPPQDSVYLVPTPSKAQQGLYQVPGSPQFQSP
PAKQTSSTFSKQTPHPFPSPATDLYQVPPGPGGPAQDIYQVPPSAGMGHDIYQVPPSMDTRSWEGTKPPA
40 KVVVPTRVGQGYVYEAQAPEQDEYDIPRHLAPGPQDIYDVPPVRGLLPSSQYQGEVYDTPPMAVKGPNGR
DPLLEVYDVPPSVEKGLPPSNHHAVYDVPPSVSKDVPDGPILLREETYDVPPAFAKAKPFDPARTPLVLAA
PPDSPPAEDVDVPPAPDLYDVPPGLRRPGPGTLYDVPRERVLPEVADGGVVDGSGVYAVPPPAEREA
PAEGKRLSASSSTRSSQASSLEVAGPGREPLELEVAVEALARLQQGVSAHVALLDLGASAGATGSW
RSPSEPQEPVLVQDLQAAVAQSAVHELLEFARSVAVGNAAHTSDRALHAKLSRQLQKMEDVHQTVAHGQ
45 ALDAGRGGSGATLEDLRLVACSRAPVEDAKQLASFLHGNASLLFRRTKATAPGPEGGGTLHPNPTDKTS
SIQSRPLPSPPKFTSQDSPDQYENSEGGWMEDYDVHLQKKEFEKTQKELLEKGSITRQKSQLELQQ
LKQFERLEQEVSRPIDHDLANWTPAQPLAPGRTGGLGPSDRQLLLFYLEQCEANLTTLTNAVDAAFTAVA
TNQPPKIFVAHSAKFVILSAHKLVFIDTSLRQAKAADVRSQVTHYSNLLCDLLRGIVATTKAAALQYPS
SAAQDMVERVKELGHSTQQFRRVLGQLAAA

50 SEQ ID NO:353

perlecan
>GL2_37_B06_G3F1 Direction: N/A
ACCGCTCCTCTCCCGAGGGCACACAACCAACCAAACTGGTGCTCTGAAACAGCGCCAGTGTCCAGTTCTCTGAGG
AGCACTGGGTCCATGAGGTCTTGGNCGGGCGGTGCAGCTGCCTGGAGCTGCTGGAGTGTGCGTGGAAGANACTG
55 TNGAGTGGCGTGGNTCATTCAAGAACNGGTGTTCCGGANGAAAACCAGAACAAGGTGGGTAGATANGGAATAAA
AGGGGGGCCGGATCGAATAAATCTTGAAGGGTCTTGAANTNAATCGTNCCAACTTTGGAACACTTGGTCCCGGG
TTTTCAACAGTGGTGGGCGAACATTTTTTCTAGCAGTGGGACCCGGCGGTTAAATAATTTCTATAAATATACT
GGGCCNGTAATAAATAGATTGGAAGNGTTCACGAGTAATTTTAAACACAACATTTCCAACCATTTGTGGGNGCCG
AGCGACCGAGAGTAATANTTAAGAAAAACGGGTACCCTTGGAAACCCCGTTTGGTGGGTGGAANAGTAAAAACA
60 AGACCCATTTATGAGGACGCGGTTNTATNTAACCAGCCCCAAAGAAACAGGTTTAAAAAAGTTTCATGGCGACC
CAGCTTNTTGGTTCGAAGTTNTGAAGACAAAGAATTTTNGGACCGNGCATTAATTTATTAATTGAGGACGGA

AAAAGAGGATTGTGTNCGGACACAGGCGTTAAAAAATTAAATTACCGCGNATAAAAGATGAAAATAGTTGGTGC
GCCGAGGNCGAAAACGCTCGGAAAGTATTCTCNGGCAGCGCCNGTTTTTATACCCGCAAAAACGAAAGAGTAG
ATAATTGTGCTAGCCGGTCGTCAANTAGATGACGAACATAAATATNATAANTATGGGGTTAACACGCTGTGGGTG
AGTAAGAGAGGAATTTAAAAAACAGGGTGTGGTTTATACATAAACATTATATTTTTATATAAGCGAATGANAGG
5 TAAAAGAGACCCGCGCNGGAGATAATATCTAGTGGGTACCAGATTGGTGAATAATAGGTACTGGGGTAAATTTGTT

SEQ ID NO:354

>GL2_37_B01_G3F1 Direction: N/A

ACCGCTCCTCTCCCGAGGGCACACACCCACCCAACTGGTGCTCTGAACCAGCGCCAGGTCCAGTTCTCTGAGGA
10 GCACTGGGTCCATGAGTCTGGNCGGCCGGTGACGCGCGGAGCTGCTGCAGGTGCTGCAGAGCCTGGAGTGGCG
TGCTCATCCAGAACTGTGTTCTGCGAGAAACAGCACAGTGGTTAGATATGATAAAGCGGGCGCTCTGACTAAAT
CTGAGGTCTTGATACTCACTGACTGTGCTAAGGGNGGAATTCTGCGGCCGCGTAAATTCAAATTTGCCCCATAN
GTGAGTCCGTATTACAATTCACCTGGGGCGTCGTTTTACAACAGTCTGTGACTGGGAAAAAACCTGGCGTTACCC
ACACATTAATCTGCCCTTGACGCGACATTCCCTTTCTGCCAGACTGGNCGTAATAGCGAAGAGGGGCCGCGACAC
15 GATCTGCCCTTCCCAACAGTTGCGCAGAGCTATACGGTACAGGGGGANGATGTTAAAGCGGNTGTTACAAACGTA
TNTAAAAGAGGTAAAGAAAAGAACCAGGAGGTGTGTAATAATNACGGAGGCNCTTANGGTGTATTGGTGCAGCGGC
GATTTGTTTNCCTAACCGCAAGNTTGGGAATACAATTTAAATTATTGAAACAAAGNGTGACNCGATGTGTGGTG
TGGAANAGGAATATAGGGGTGT

SEQ ID NO:355

>GL2_37_B02_G3F1 Direction: sense

ACCGCTCCTCTCCCGAGGGCACACACCCACCCAACTGGTGCTCTGAACCAGCGCCAGGTCCAGTTCTCTGAGG
AGCACTGGGTCCATGAGTCTGGCCGGCCGGTGACGCGCGGAGCTGCTGCAGGTGCTGCAGAGCCTGGAGGCCG
25 TGCTCATCCAGACCGTGT

SEQ ID NO:356

>GL2-86-1M13F Direction: sense

CTGGTGGGGATGGGGACCGCTGCCAGGGGTGAGCTGCCTTTTGCTCCACAGCCGACACTAAAGACAATTCCCA
ATCCTGAGTGGGTGGCAGAGACTCCTGCGATGCCCGTCTCAGGTAGCTGTGGGGCACCAGCCACAAGCCGAGGT
30 TGGCTCTCCTAGGAGTGAGAACTGCCCAAGGGCTGCAGAAACAGGCCACCCAGCTCTATCTGGGGGCTCCATCGG
TGGGTAGGGGGACAGTGGGGGCGAGTTCTGGGGCCACCCAGCCACTGTTCTGACCCCAAGTCTTGGTGACTTTCT
GAGGTGCCCACTCCCATCCAACCTGCCTTGCTGGCCAGCCTTGTGGCTTTGCCAGCTGTGTGTGTGAGGGTGGC
ATGCCACCTCCAGTCCAGCCAGGGCGGTA

SEQ ID NO:357

>GL2-86-2M13F Direction: sense

CTGGTGGGGATGGGGACCGCTGCCAGGGGTGAGCTGCCTTTTGCTCCACAGCCGGCACTAAAGACAATTCCCA
TCCTGAGTGGGTGGCAGAGACTCCTGCGATGCCCGTCTCANGTAGCTGTGGGGCACCAGCCACAAGCCGAGGTT
GGNTCTCCTAGGAGTGAGAACTGCCCAAGGGCTGCAGAAACAGGCCACCCAGCTCTATCTGGGGGCTCCAATCGG
40 TGGGTAGGGGGACAGTGGGGGCGAGTTCTGGGGCCACCCAGCACTGTTCTGAACCCAAGTCTTGGTGACTTTCTG
AGGTGCCCACTCCCATCCAACCTGGCTTGCTGGNAGCCTTGTGGCTTTGNCCAGCTGTGTGTGTGAGGGTGGCA
TGCCCANCTCCAGTCCAGCCAGGGCGGTAGCAGCAAAGCGTGGCATCGACTCGGNNTTTCTACAAAAAATTCA
TAAATAATATTCACTAATAATACTCGGAACATTTGTGGGGCTTGGGGGCGTTGGCCCCGGGGAAAGTCCAGTG
TTGGGGGGCAAGGCCAGGTTGGCCCTANACGAAGGGGGGCCAGGGGCCGTTGTGTTTGGGGCCCCGGGCCCTGGGG
45 CGCNGGTGCTTGACAGGGC

SEQ ID NO:358

>GL2-86-3M13F Direction: sense

CTGGTGGGGATGGGGACCGCTGCCAGGGGTGAGCTGCCTTTTGCTCCACAGCCGGCACTAAAGACAATTCCCA
50 ATCCTGAGTGGGTGGCAGAGACTCCTGCGATGCCCGTCTCAGGTAGCTGTGGGGCACCAGCCACAAGCCGAGGT
TGGCTCTCCTAGGAGTGAGAACTGCCCAAGGGCTGCAGAAACAGGCCACCCAGCTCTATCTGGGGGCTCCATCGG
TGGGTAGGGGGACAGTGGGGGCGAGTTCTGGGGCCACCCAGCCACTGTTCTGACCCCAAGTCTTGGTGACTTTCT
GAGGTGCCCACTCCCATCCAACCTGCCTTGCTGGCCAGCCTTGTGGCTTTGCCAGCTGTGTGTGTGAGGGTGGC
ATGCCACCTCCAGTCCAGCCAGGGCGGTAGCAGCAAAGCGTGGCATCGCCTCGGTTTCTTACAAAAAATTCATA
55 ATAATATTAATAATAATACTCGACATTGTGCGGGCTG

SEQ ID NO:359

>GL2-86-5M13F Direction: sense

CTGGTGGGGATGGGGACCGCTGCCAGGGGTGAGCTGCCTTTTGCTCCACAGCCGGCACTAAAGACAATTCCCA
60 ATCCTGAGTGGGTGGCAGAGACTCCTGCGATGCCCGTCTCAGGTAGCTGTGGGGCACCAGCCACAAGCCGAGGT

TGGCTCTCCTAGGAGTGAGAACTGCCCAAGGGCTGCAGAAACAGGCCACCCAGCTCTATCTGGGGGCTCCATCGG
TGGGTAGGGGGACAGTGGGGGCAGTTCTGGGCCCACCCAGCCACTGTNCCTGACCCCAAGTCTGGTGACTTTCT
GAGGTGCCCACTCCACATCCAACCTGCCTTGCTGGNCAGCCTTGNTGGCATAATGCCCAGNCTGATAGATGCGTG
AGGGNGGACATGCCACCTACCAGTACCAGACCCAGGGCGGATAGGCAGCAAAGNACGTGGCATCGCCTCGGG

SEQ ID NO:360

>GL2-86-1M13R Direction: N/A

CTGGTCAGCGGCCGGTCCCCAGGTCCCAACGTGGCAGTCAACGCCAAGGGCAGCGTCTACATCGGCGGAGCCCCCT
GACGTGGCCACGCTGACCGGGGGCAGATTCTCCTCGGGCATCACAGGCTGTGTCAAGAACCTGGTGCTGCACTCG
GCCCCACCCGGCGCCCCGCCCCACAGCCCCTGGACCTGCAGACCGGCCAGGCCGGGGGCAACACACGCCCCCTG
CCCCTCGTAGGCACCTGCCTGCCCCACACGACTCCCGGGCCACGCCCCAGCCCGACAATGTCGAGTATATTATT
ATTAATATTATTATGAA

SEQ ID NO:361

>GL2-86-2M13R Direction: N/A

CTGGTCAGCGGCCGGTCCCCAGGTCCCAACGTGGCAGTCAACGCCAAGGGCAGCGTCTACATCGGCGGAGCCCCCT
GACGTGGCCACGCTGACCGGGGGCAGATTCTCCTCGGGCATCACAGGCTGTGTCAAGAACCTGGTGCTGCACTCG
GCCCCACCCGGCGCCCCGCCCCACAGCCCCTGGACCTGCAGACCGGCCAGGCCGGGGGCAACACACGCCCCCTG

SEQ ID NO:362

>GL2-86-3M13R Direction: N/A

TTGGTCAGCGGCCGGTCCCCAGGTCCCAACGTGGCAGTCAACGCCAAGGGCAGCGTCTACATCGGCGGAGCCCCCT
GACGTGGCCACGCTGACCGGGGGCAGATTCTCCTCGGGCATCACAGGCTGTGTCAAGAACCTGGTGCTGCACTCG
GCCCCACCCGGCGCCCCGCCCCACAGCCCCTGGACCTGCAGACCGCGCCCAGGCCGGGGGCAACACACGCCCCCT
TGCCCCCTCGTAGGCACCTGCCTGCCCCACACGACTCCCGGGCCACGCCCCAGCCCGACAATGTCGAGTATATTATT
TTATTAATATTATTATGAATTTTNNGTAAAGAAACCGAGGCGATGCCACGCTTANGCTGCTACCGCCCTGGGCTGG
ACTGGAGGTGGGCATGCCACC

SEQ ID NO:363

>GL2-88-5M13R Direction: anti-sense

CTGGTCAGCGGCCGGTCCCCAGGTCCCAACGTGGCAGTCAACGCCAAGGGCAGCGTCTACATCGGCGGAGCCCCCT
GACGTGGCCACGCTGACCGGGGGCAGATTCTCCTCGGGCATCACAGGCTGTGTCAAGAACCTGGTGCTGCACTCG
GCCCCACCCGGCGCCCCGCCCCACAGCCCCTGGACCTGCAGACCGCGCCCAGGCCGGGGGCAACACACGCCCCCT
GCCCCCTCGTAGGCACCTGCCTGCCCCACACGACTCCCGGGCCACGCCCCAGCCGACAATGTCGAGTATATTATT
ATTAATATTATTATGAATTTTGTAAAGAAACCGAGGCGATGCCACGCTTTGCTGCTACCGGCCTGGGGCTGGA
CTGGAGGTGGGCATGCCAACACCATCACGGCACAACAGCTGGCAAAGCCAAAGAGGGCCTTGGGGCCAGGCAAG
GGCACAGGTTGGGACTGGGAGTGGGGCACCCTCANGAAAGGTCACCCAGGGAATATGGGGGGGTNCCAGAGGAA
ACCACAGGTAGGGGACTGGGGGTTTNGGGGGCCCCCAGGAAACTTGGGGCGCCCAAAANATNGNTACCACC
ACCATAAACACACAAGCCCAGAANTGTGAAAACCCCCAACCCAGAGAAAANTAGCAGAGCCATCGAGGAAAGCACC
CCAAGATACACAGCANNACCCCATGAGAA

SEQ ID NO:364

>GL2_93_2B09_G3F1 Direction: N/A

AAGCTGCCTTTTGCTCCACAGCCGGCACTAAAGACAATTCCTGAGTGGGTGGCAGAGACTCCCTTTGGG
GCGATGCCCGTCTTCCAGGTAAGGCTGTGGTGGGCGGGCCAACACACCAAAAAGGGGCCCCCCCCCAAAATAGCC
CCATACAAACAAAACACGNNGGGGCCATACCCACGGGAAAAGGGGGGTTTATGGGTGGGGCCATTTTCCCCCAGT
TTCCACCCGTTTAAACAGGCGNNGGCAACACNGCTTTTGGGCGTGNCTGGAACAACAACCTTTTGGGTCTCCTCC
NCCCAAAAAGTAGGGTGGGCTGGGCGGGCCCCCTTTATGTGTGGCCCAANACNGGAGAACACAATACTAAATAATC
CCAAACTGGTGTGGGGTGTCCCGCCCCAAATANACCCACCCCAANTAGATGCCCTTTTACCCCGTTTTATAAAA
AATTTTTTTTCCCCCTATATCTTGGCGGGGGCGAGGGGATATGT

SEQ ID NO:365

>gi|7427516|ref|NM_005529.2| Homo sapiens heparan sulfate proteoglycan 2 (perlecan) (HSPG2), mRNA

GCCCCGAGCGAGCGAGCGAGAGAGCGGCGCGGGCCCGGGCCATGGGGTGGCGGGCGCCGGGCGCGCTGCTG
CTGGCGCTGCTGCTGCACGGGCGGCTGCTGGCGGTGACCCATGGGCTGAGGGCATAAGATGGCTTGTCTC
TGCTTGAGGACATAGAGACCGTCACAGCAAGCCAAATGCGCTGGACACATTCTGTACCTTTTCTGATGATGA
GGACATGCTGGCTGACAGCATCTCAGGAGACGACCTGGGCAGTGGGGACCTGGGCAGCGGGGACTTCCAG
ATGTTTATTTCCGAGCCCTGGTGAATTTCACTCGCTCCATCGAGTACAGCCCTCAGCTGGAGGATGCAG
GCTCCAGAGAGTTTCGAGAGGTGTCCGAGGCTGTGGTAGACACGCTGGAGTCGGAGTACTTGAAAATTCC

CGGAGACCAGGTTGTCACTGTGGTGTTCATCAAGGAGCTGGATGGCTGGGTTTTTGTGGAGCTGGATGTG
GGCTCGGAAGGGAATGCGGATGGGGCTCAGATTAGGAGATGCTGCTCAGGGTCATCTCCAGCGGCTCTG
TGGCCTCCTACGTACCTCTCCCCAGGGATTCCAGTTCCGACGCCTGGGCACAGTGCCCCAGTTCCCAAG
AGCCTGCACGGAGGCCGAGTTTGCCTGCCACAGCTACAATGAGTGTGTGGCCCTGGAGTATCGCTGTGAC
5 CGGCGGGCCCGACTGCAGGGACATGTCTGATGAGCTCAATTGTGAGGAGCCAGTCTTGGGTATCAGCCCCA
CATTTCTCTCTCTCGTGGAGACGACATCTTTACCGCCCCGGCCAGAGACAACCATCATGCGACAGCCACC
AGTACCCACGCTCCTCAGCCCCCTGCTTCCCGGTTCCGTCAGGCCCCCTGCCCTGTGGGCCCCAGGAGGCC
GCATGCCGCAATGGGCACTGCATCCCCAGAGACTACCTCTGCGACGGACAGGAGGACTGCGAGGACGGCA
GCGATGAGCTAGACTGTGGCCCCCGCCACCCTGTGAGCCCCAACGAGTTCCCCCTGCGGGAATGGACATTG
10 TGCCCTCAAGCTGTGGCGCTGCGATGGTGACTTTGACTGTGAGGACCGAAGTGTGAAGCCAAGTGGCCCC
ACCAAGCGTCTGAGGAAGTGTGCGGGCCCCACACAGTTCGGATGCGTCTCTACCAACATGTGCATCCAG
CCAGCTTCCACTGTGACGAGGAGAGCGACTGTCTGACCGGAGCGACGAGTTTGGCTGCATGCCCCCCCA
GGTGGTGACACCTCCCCGGGAGTCCATCCAGGCTTCCCGGGGCCAGACAGTGACCTTCACCTGCGTGGCC
ATTGGCGTCCCCGCCCCCTTTCTCATCAATTGGAGGCTCAACTGGGGCCACATCCCCCTCTCAGCCCAGGG
15 TGACAGTGACCAGCGAGGGTGGCCGTGGCACACTGATCATCCGTGATGTGAAGGAGTCAGACCAGGGTGC
CTACACCTGTGAGGCCATGAACGCCCCGGGGCATGGTGTGTTGGCATTCTTGACGGTGTCTTTGAGCTCGTC
CCACAACGAGCAGGCCCCCTGCCCTGACGGCCACTTCTACCTGGAGCACAGCGCCGCTGCCTGCCCTGCT
TCTGCTTTGGCATCACCAGCGTGTGCCAGAGCACCCGCGCTTCCGGGACCAGATCAGGCTGCGCTTTGA
CCAACCCGATGACTTCAAGGGTGTGAATGTGACAATGCCTGCGCAGCCCCGGCACGCCACCCCTCTCCTCC
20 ACGCAGCTGCAGATCGACCCATCCCTGCACGAGTTCCAGCTAGTCGACCTGTCCCGCCGCTTCTCTGTC
ACGACTCCTTCTGGGCTCTGCCTGAACAGTTCCTGGGCAACAAGGTGGACTCCTATGGCGGCTCCCTGCG
TTACAACGTGCGCTACGAGTTGGCCGTGGCATGCTGGAGCCAGTGACGCGCCGACGCTGGTCTCTGTCG
GGTCCGGGTACCGCCTCTCTCCGAGGCCACACACCCACCCAACCTGGTGCTCTGAACCAGCGCCAGG
TCCAGTTCTCTGAGGAGCACTGGGTCCATGAGTCTGGCCGGCCGGTGCAGCGCGCGGAGCTGCTGCAGGT
25 GCTGCAGAGCCTGGAGGCCGTGCTCATCCAGACCGTGTACAACACCAAGATGGCTAGCGTGGGACTTAGC
GACATCGCCATGGATACCACCGTCACCCATGCCACCAGCCATGGCCGTGCCACAGTGTGGAGGAGTGCA
GATGCCCCATTGGCTATTCTGGCTTGTCTGCGAGAGCTGTGATGCCCACTTCACTCGGGTGCCTGGTGG
GCCCTACCTGGGCACCTGCTCTGGTTGCAGTTGCAATGGCCATGCCAGCTCCTGTGACCTGTGTATGGC
CACTGCCTGAATTGCCAGCACAAACCGGAGGGGCCACAGTGCAAGAAAGTGCAAGGCTGGCTTCTTTGGGG
30 ACGCCATGAAGGCCACGGCCACTTCTGCGCGCCCTGCCCTTGCCCATACATCGATGCCCTCCCGCAGATT
CTCAGACACTTGCTTCTGACACGATGGCCAGCCACATGTGACGCTGTGCCCCAGGCTACACTGGC
CGCCGCTGTGAGAGCTGTGCCCCCGGATACGAGGGCAACCCCATCCAGCCCGGCGGGAAGTGACGGCCCG
TCAACCAGGAGATTGTGCGCTGTGACGAGCGTGGCAGCATGGGGACCTCCGGGGAGGCCTGCCGCTGTAA
GAACAATGTGGTGGGGCGCTTGTGCAATGAATGTGCTGACCGCTCTTTCCACCTGAGTACCCGAAACCCC
35 GATGGCTGCCTCAAGTGCTTCTGCATGGGTGTGAGTCCGCACTGCACCAGCTCTTCATGGAGCCGTGCCC
AGTTGCATGGGGCCTCTGAGGAGCCTGGTCACTTCAGCCTGACCAACGCGCAAGCACCCACACCACAA
CGAGGGCATCTTCTCCCCACGCCCCGGGGAACCTGGGATTCTCCTCCTTCCACAGACTCTTATCTGGACCC
TACTTCTGGAGCCTCCCTTACGCTTCTCTGGGGGACAAGGTGACCTCCTATGGAGGAGAGCTGCGCTTCA
CAGTGACCCAGAGGTCCAGCCGGGCTCCACACCCCTGCACGGGCAGCCGTTGGTGGTGCTGCAAGGTTAA
40 CAACATCATCCTAGAGCACCATGTGGCCACAGGACCCAGCCCCGGCCAGCCAGCACCTTCATTGTGCCT
TTCCGGGAGCAAGCATGGCAGCGGCCCGATGGGCAGCCAGCCACACGGGAGCACCTGCTGATGGCACTGG
CAGGCATCGACACCTCCTGATCCGAGCATCTACGCCAGCAGCCCGCTGAGAGCAGGGTCTCTGGCAT
CAGCATGGACGTGGCTGTGCCCGAGGAAACCGGCCAGGACCCCGCGCTGGAAGTGGAACAGTGCTCCTGC
CCACCCGGGTACCGTGGGCGCTCCTGCCAGGACTGTGACACAGGCTACACACGCACGCCCAGTGGCCTCT
45 ACCTGGGTACCTGTGAACGCTGCAGCTGCCATGGCCACTCAGAGGCCCTGCGAGCCAGAAACAGGTGCCTG
CCAGGGCTGCCAGCATCACACGGAGGGCCCCCGGTGTGAGCAGTGCCAGCCAGGATACTACGGGGACGCC
CAGCGGGGGACACCACAGGACTGCCAGCTGTGCCCTGCTACGGAGACCTGCTGCGCGGAGGCTGCCCC
ACACTTGTTTTCTGGACACAGACGGCCACCCACCTGTGATGCGTGCTCCCCAGGCCACAGTGGGCGTCA
CTGTGAGAGGTGCGCCCCCTGGCTACTATGGCAACCCAGCCAGGGCCAGCCATGCCAGAGAGACAGCCAG
50 GTGCCAGGGCCCCATAGGCTGCAACTGTGACCCCCAAGGCAGCGTCAGCAGCCAGTGTGATGCTGCTGGTC
AGTGCCAGTGCAAGGCCAGGTGGAAGGCCTCACTTGCAGCCACTGCCGGCCCCACCACTTCCACCTGAG
TGCCAGCAACCCAGACGGCTGCCTGCCCTGCTTCTGTATGGGCATCACCCAGCAGTGCGCCAGCTCTGCC
TACACACGCCACCTGATCTCCACCCACTTTGCCCTGGGGACTTCCAAGGCTTTGCCCTGGTGAACCCAC
AGCGAAACAGCCGCTGACAGGAGAAATTCATGTGAACCCGTGCCCGAGGGTGGCCAGCTCTCTTTTG
55 CAACTTTGCCAACTCGGCCATGAGTCTTCTACTGGCAGCTGCCGGAGACATAACAGGGAGACAAGGTG
GCGGCCTACGGTGGGAAGTTGCGATACACCTCTCTACAGCAGGCCACAGGGCAGCCCACTCTCTG
ACCCCGATGTGCAGATCACGGGCAACAACATCATGCTAGTGGCCTCCAGCCAGCGCTGCAGGGCCCTGA
GAGGAGGAGCTACGAGATCATGTTCCGAGAGGAATTCGGCGCCGGCCCCGATGGGCAGCCGGCCACACGC
GAGCACCTCCTGATGGCACTGGCCGACCTGGATGAGCTCCTGATCCGGGCCACGTTCTCCTCCGTGCCGC
60 TGGTGGCCAGCATCAGCGCAGTCAGCCTGGAGGTGCCCCAGCCGGGGCCCTCAAACAGACCCCGCGCCCT
CGAGGTGGAGGAGTGCCGCTGCCCGCCAGGCTACATCGGTCTGTCTGCTGCCAGGACTGTGCCCCCGGCTAC

ACGCGCACCGGGAGTGGGCTCTACCTCGGCCACTGCGAGCTATGTGAATGCAATGGCCACTCAGACCTGT
GCCACCCAGAGACTGGGGCTTGCTCGCAATGCCAGCACAACGCCGAGGGGAGTTCTGCGAGCTTTGTGC
CCCTGGCTACTACGAGATGCCACAGCCGGGACGCCTGAGGACTGCCAGCCCTGTGCCCTGCCCACTGACC
5 AACCAGAGAACATGTTTCCCGCACCTGTGAGAGCCTGGGAGCCGGCGGGTACCGCTGCACGGCCTGCG
GGGCCAGTGCCCTGCCAGAGACAAACAGCCCCACTGGTGGTTCGAGGTCCATCCTGCTCGAAGCATAGTG
CCCCAAGGTGGCTCCCACTCCCTGCGGTGTGAGGTGAGTGGGAGGGGACCCCACTACTTCTATTGGTCCC
GTGAGGATGGGCGCCTGTGCCAGCGGCACCCAGCAGCGACATCAAGGCTCCGAGCTCCACTTCCCCAG
10 CGTCCAGCCCTCGGATGCTGGGGTCTACATTTGCACCTGCCGTAATCTCCACCGATCCAATACCAGCCGG
GCAGAGCTGCTGGTCACTGAGGCTCCAAGCAAGCCCATCACAGTGAAGTGTGGAGGAGCAGCGGAGCCAGA
GCGTGCGCCCCGGAGCTGACGTACCTTTCATCTGCACAGCCAAAAGCAAGTCCCCAGCCTATACCCTGGT
GTGGACCCCGCTGCACAACGGGAAACTGCCACCCGAGCCATGGATTTCAATGGCATCCTGACCATTTCGC
AACGTCCAGCTGAGTGATGCAGGCACCTACGTGTGCACCGGCTCCAACATGTTTGGCATGGACCAGGGCA
15 CAGCCACTCTACATGTGTCAGGCTCGGGCACCTTGTCCGCCCCCGTGGTCTCCATCCATCCGCCACAGCT
CACAGTGACGCCCCGGGCAACTGGCGGAGTTCCGCTGCAGCGCCACAGGGAGCCCCACGCCACCCTCGAG
TGGACAGGGGGCCCCGGCGCCAGCTCCCTGCGAAGGCACAAATCCACGGCGGCATCCTGCGCCTGCCAG
CTGTGAGCCCCACGGATCAGGCCCAGTACTTGTGCCGAGCCCCACAGCAGCGCTGGGCAGCAGGTGCCAG
GGCTGTGCTCCACGTGCATGGGGCGGTGGGCCCAGAGTCCAAGTGAGCCAGAGAGGACCCAGGTCCAC
20 GCAGGCCGACCCGTGAGGCTGTACTGCAGGGCTGCAGGCGTGCCTAGCGCCACCATCACCTGGAGGAAGG
AAGGGGGCAGCCTCCACCACAGGCCCCGTGAGAGCGCACAGACATCGCGACACTGCTCATCCCAGCCAT
CACGACTGCTGACGCCGGCTTCTACCTCTGCGTGGCCACCAGCCCTGCAGGCACTGCCAGGCCCCGGATC
CAAGTGGTTGCTCTTTCAGCCTCAGATGCCAGCCAACCGCCGGTCAAGATTGAGTCTCATCGCCTTCTG
TGACAGAAGGGCAAACTCGACCTCAACTGTGTGGTGGCAGGGTCAGCCCATGCCAGGTACCTGGTA
25 CAGGCGAGGGGGTAGCCTGCCTCACCACACCCAGTGCACGGCTCCCGTCTGCGGCTCCCCAGGTCTCA
CCAGCTGATTCTGGAGAGTATGTGTGCCGTGTGGAGAATGGATCGGGCCCCAAGGAGGCCCTCCATTACTG
TGTCTGTGCTCCACGGCACCCATTCTGGCCCCAGCTACACCCAGTGCCCGGCAGCACCCGGCCCCATCCG
CATCGAGCCCTCCTCCTCACACGTGGCGGAAGGGCAGACCCTGGATCTGAAGTGCCTGGTGGCCGGCAG
GCCCCAGCCAGGTACGTGGCACAAGCGTGGGGGCAGCCTCCCTGCCCGGCACCAGACCCACGGCTCGC
30 TGCTGCGGCTGCACCAGGTGACCCCGGCGGACTCAGGCGAGTATGTGTGCCATGTGGTGGGCACCTCCGG
CCCCCTAGAGGCCTCAGTCTGGTCAACATCGAAGCCTCTGTCTATCCCTGGACCCATCCCACCTGTGAGG
ATCGAGTCTTTCATCTCCACAGTGGCCGAGGGCCAGACCCTGGATCTGAGCTGCGTGGTGGCAGGGCAGG
CCACGCCCCAGGTACATGGTACAAGCTGCGGGCAGCCTCCCTGCCCGGCACCAGGTTGCTGGCTCCCG
CCTGTACATCTTCCAGGCCCTCACCTGCCGATGCGGGACAGTACGTCTGCCGGGCCAGCAACGGCATGGAG
35 GCCTCCATCACGGTCACAGTAAGTGGGACCCAGGGGGCCAACTTAGCCTACCCTGCCGGCAGCACCCAGC
CCATCCGCATCGAGCCCTCCTCCTCGCAAGTGGCGGAAGGGCAGACCCTGGATCTGAAGTGCCTGGTGCC
CGGGCAGTCCCATGCCAGGTACGTGGCACAAGCGTGGGGGCAGCCTCCCTGTCCGGCACCAGACCCAC
GGCTCCCTGCTGAGACTCTACCAAGCGTCCCCCGCCGACTCGGGCGAGTACGTGTGCCGAGTGTGGGCA
GCTCCGTGCCCTTAGAGGCCTCTGTCTGGTCAACATTGAGCCTGCGGGCTCAGTGCCTGCACTTGGGGT
40 CACCCCCACGGTCCGGATCGAGTCACTGCTTTCGCAAGTGGCAGAGGGGCAGACCCTGGACCTGAAGTGC
CTCGTTGCTGGTCCAGGCCATGCCAGGTACGTGGCACAAGCGCGGGGGCAGCCTCCCGGCCGGCACC
AGGTGCATGGCTCAGGCTACGCGTCTCCAGGTGACCCAGCTGATTACAGGGAGTACGTGTGCCGTGT
GGTCCGCAGCTCAGGTACCCAGGAAGCCTCAGTCTTGTGTCACCATCCAGCAGCGCCTTAGTGGCTCCCAC
TCCCAGGGTGTGGCGTACCCCGTCCGCATCGAGTCTCCTCAGCCTCCCTGGCCAATGGACACACCCTGG
45 ACCTCAACTGCCTGGTTGCCAGCCAGGCTCCCCACACCATCACCTGGTATAAGCGTGGAGGCAGCTTACC
CAGCCGGCACCAGATCGTGGGCTCCCGGCTGCGGATCCCTCAGGTGACTCCGGCAGACTCGGGCGAGTAC
GTGTGTACGTCAGTAACGGTGCAGGCTCCCGGGAGACCTCGCTCATCGTCACCATCCAGGGCAGCGGTT
CCTCCCACGTGCCAGAGTCTCCCCACCGATCAGGATCGAGTCTCTTCCCCACGGTGGTGGAAAGGGCA
GACCTTGGATCTGAAGTGCCTGGTTCGCCAGGCAGCCCCAGGCTATCATCACATGGTACAAGCGTGGGGGC
50 AGCCTTCCCTCCCCGACACCAGACCCATGGCTCCACCTGCGGTTGCACCAAATGTCTGTGGCTGACTCGG
GCGAGTATGTGTGCCGGGCCAACAACAACATCGATGCCCTGGAGGCTCAATCGTCTATCTCCGTCTCCCC
TAGCGCCGGCAGCCCCCTCCGCCCTGGCAGCTCCATGCCCATCAGAATTGAGTCACTCTCTCACAGTG
GCCGAAGGGGAGACCCCTGGATCTGAAGTGCCTGGTCCCGGGCAGGCCCATGCCAGGTCACTTGGCACA
AGCGTGGGGGAGCCTCCCACTTACCATCAGACCCGCGGCTCACGGCTGCGGCTGCACCATGTGTCCCC
GGCCGACTCGGGTGAATACGTGTGCCGGGTGATGGGCAGCTCTGGCCCCCTGGAGGCCTCAGTCTCTGTC
55 ACCATCGAAGCCTCTGGCTCAAGTGTGTCCACGTCCCCGCCCCAGGTGGAGCCCCACCCATCCGCATCG
AGCCCTCCTCCTCCGAGTGGCAGAAGGGCAGACCCTGGATCTGAAGTGCCTGGTGGTGGCCGGCAGGCCCA
CGCCAGGTACGTGGCACAAGCGTGGAGGAAACCTCCCTGCCCGGCACCAGGTCCACGGGCCACTGCTG
AGGTGAACAGGTGTCCCCGGCTGACTCTGGCGAGTACTCGTGCCAAGTGACCGGAAGCTCAGGCACCC
TGGAGGCATCTGTCTGGTCAACAATTGAGCCTCCAGCCAGGACCCATTCTGCTCCAGGACTGGCCCCA
60 GCCATCTACATCGAGGCTCCTCTTTCACAGTGAAGTGAAGGGCAGACTCTGGATCTGAAGTGTGTGGT
CCCGGGCAGGCCCATGCCAGGTACGTGGTACAAGCGCGGGGGCAGCCTCCCCCGCCCGCACCAGACCC

ATGGCTCCAGCTGCGGCTCCACCACGTCTCCCCTGCCGACTCAGGCGAGTATGTGTGTCGTGCAGCCGG
CGGCCCAGGCCCTGAGCAAGAAGCCTCCTTCACAGTCACCGTCCCGCCAGTGAGGGGTCTTCTTACCGC
CTTAGGAGCCCGGTCTATCTCGATCGACCCGCCAGCAGCACCGTGCAGCAGGGCCAGGATGCCAGCTTCA
AGTGCCTCATCCATGACGGGGCAGCCCCCATCAGCCTCGAGTGAAGACCCGGAACCAGGAGCTGGAGGA
5 CAACGTCCACATCAGTCCCAATGGCTCCATCATCACCATCGTGGGCACCCGGCCAGCAACCACGGTACC
TACCGCTGCGTGGCCTCCAATGCCTACGGTGTGGCCAGAGTGTGGTGAACCTCAGTGTGCACGGGCCCC
CTACAGTGTCCGTGCTCCCCGAGGGCCCCGTGTGGGTGAAAGTGGGAAAGGCCGTACCCCTGGAGTGTGT
CAGTGCCGGGGAGCCCCGCTCCTCTGCTCGTTGGACCCGGATCAGCAGCACCCCTGCCAAGTTGGAGCAG
CGGACATATGGGCTCATGGACAGCCACACGGTGTGTCAGATTTCATCAGCTAAACCATCAGATGCGGGCA
10 CTTATGTGTGCCCTTGCTCAGAATGCACTAGGCACAGCACAGAAGCAGGTGGAGGTGATCGTGGACACGGG
CGCCATGGCCCCAGGGGCCCTCAGGTCCAAGCTGAAGAAGCTGAGCTGACTGTGGAGGTGGACACAG
GCCACCTTGCGCTGCTCAGCCACAGGCAGCCCCGCGCGCACCATCCACTGGTCCAAGCTGCGTTCCCCAC
TGCCCTGGCAGCACCGGCTGGAAGGTGACACACTCATCATACCCCGGGTAGCCAGCAGGACTCGGGCCA
GTACATCTGCAATGCCACTAGCCCTGCCGGGCACGCTGAGGCCACCATCATCTGCACGTGGAGAGCCCCA
15 CCATATGCCACCACGGTCCCAGAGCAGCTTCGGTGCAGGCAGGGGAGACGGTGCAGCTCCAGTGCCTGG
CTCACGGGACACCCCCACTCACCTTCCAGTGGAGCCGCGTGGGCAGCAGCCTTCTTGGGAGGGCGACCGC
CAGGAACGAGCTGCTGCACTTTGAGCGTGCAGCCCCCTGAGGACTCAGGCCGCTACCGCTGCCGGGTACCC
AACAAGGTGGGCTCAGCCGAGGCCTTTGCCAGCTGCTCGTCCAAGGCCCTCCCGGCTCTCTCCCTGCCA
CCTCCATCCCAGCAGGGTCCACGCCACCGTGCAGGTACGCCCTCAGCTAGAGACCAAGAGCATTGGGGC
20 CAGCGTTGAGTTCCATTGTGCTGTGCCAGCGACGGGTACCCAGCTCCGTTGGTTCAAGGAAGGGGGT
CAGCTGCCTCCCGGTACAGCGTGCAGGATGGGTGCTCCGAATCCAGAACTGGACCAGAGTGGCCAAAG
GGACGTATATATGCCAGGCCCATGGACCTTGGGGGAAGGCCAGGCCAGTGCCAGCTGGTTATCCAAGC
CCTGCCCTCGGTGCTCATCAACATCCGGACCTCTGTGCAGACCGTGGTGGTTGGCCACGCCGTGGAGTTC
GAATGCCTGGCACTGGGTGACCCCAAGCCTCAGGTGACATGGAGCAAAGTTGGAGGGCACCTGCGGCCAG
25 GCATTGTGCAGAGCGGAGGTGTCTGTCAGGATCGCCACGTAGAGCTGGCTGATGCGGGACAGTATCGCTG
CACTGCCACCAACGCAGCTGGCACCACACAATCCCACGTCTGTGCTGCTTGTGCAAGCCTTGCCCCAGATC
TCAATGCCCCAAGAAGTCCGTGTGCTGCTGGTTCTGCAGCTGTCTTCCCCCTGCATAGCCTCAGGCTACC
CCACTCTGACATCAGCTGGAGCAAGCTGGATGGCAGCCTGCCACCTGACAGCCGCTGGAGAACAACT
GCTGATGCTGCCCTCAGTCCCAACCCAGGACGCGAGTACCTACGTCTGCACCGCCACTAACCCGACAGGC
30 AAGGTCAAAGCCTTTGCCACCTGCAGGTGCCAGAGCGGGTGGTGGCCCTACTTCACGCAGACCCCCCTACT
CCTTCTTACCGCTGCCCCACCATCAAGGATGCCTACAGGAAGTTCGAGATCAAGATCACCTTCCGGCCTGA
CTCAGCCGATGGGATGCTGCTGTACAATGGGCAGAAGCGAGTCCCAGGGAGCCCCACCAACCTGGCCAAC
CGGCAGCCCGACTTCATCTCCTTCGGCCTCGTGGGGGGAAGGCCGAGTTCCGGTTCGATGCAGGCTCAG
GCATGGCCACCATCCGCCATCCACACCACTGGCCCTGGGCCATTTCCACACCGTGACCCTGTGCGCAG
35 CCTCACCCAGGGCTCCCTGATTGTGGGTGACCTGGCCCCGGTCAACGGGACCTCCCAGGGCAAGTTCCAG
GGCCTGGATCTGAACGAGGAACCTTACCTGGGTGGCTATCCTGACTATGGTGCCATCCCCAAGGCGGGGC
TGAGACAGCGGCTTCATAGCTGTGTCCGGGAGCTGCGCATCCAGGGCGAGGAGATCGTCTTCCATGACCT
CAACCTCACGGCGCACGGCATCTCCCACTGCCCCACCTGTGCGGACCGGCCCTGCCAGAATGGCGGTGAG
TGCCATGACTCTGAGAGCAGCAGCTATGTGTGCGTCTGCCAGCTGGCTTACCAGGGAGCCGCTGTGAGC
40 ACTCGCAGGCCCTGCACTGCCATCCAGAGGCCTGTGGGGCCGACGCCACCTGTGTGAACCGGCCTGACGG
TCGAGGCTACACCTGCCGCTGCCACCTGGGCCGCTCGGGGTGTCGGGTGTGAGGAAGGTGTGACAGTGACC
ACCCCTCGCTGTGCGGTGCTGGCTCCTACCTGGCACTGCCCGCCCTACCAACACACACCACGAGCTAC
GCCTGGACGTGGAGTTCAAGCCACTCGCCCCCTGACGGGTCTCTGCTGTTTCAGCGGGGGGAAGAGCGGGCC
TGTGGAGGACTTCGTGTCCCTGGCGATGGTGGGCGGCCACCTGGAGTTCCGCTATGAGTTGGGGTCAGGG
45 CTGGCCGTTCTGCGGACGGCCGAGCCGCTGGCCCTGGGCCGCTGGCACCGTGTGCTGTCAGAGCGTCTCA
ACAAGGACGGCAGCCTGCGGGTGAATGTTGGACGCCCTGTGCTGCGCTCCTCGCCCGCAAGAGCCAGGG
CCTCAACCTGCACACCCTGCTCTACCTGGGGGGTGTGGAGCCTTCCGTGCCACTGTCCCCGGGCCACCAAC
ATGAGCGCTCACTTCCGCGGCTGTGTGGGCGAGGTGTCAGTGAATGGCAAACGGCTGGACCTCACCTACA
GTTTCTTAGGCAGCCAGGGCATCGGGCAATGCTATGACAGCTCCCCATGTGAGCGCCAGCCTTGCCAACA
50 TGGTGCCACGTGCATGCCCGCTGGCGAGTATGAGTTCCAGTGCCTGTGTGTCGAGATGGAATCAAAGGAGAC
CTGTGTGAGCAGGAGAGAACCCTGCCAGCTCCGTGAACCTGTCTGCATGGGGGCACCTGCCAGGGCA
CCCGCTGCCTCTGCCTCCCTGGCTTCTCTGGCCCCACGCTGCCAACAAGGCTCTGGACATGGCATAGCAGA
GTCCGACTGGCATCTTGAAGGCAGCGGGGGCAATGATGCCCTGGGACGTACGGAGCCTATTTCCACGAT
GATGGCTTCTCTGCCTTCCCTGGCCATGTCTTCTCCAGGAGCCTGCCCCAGGTGCCCGAGACCATCGAGC
55 TGGAGGTTTCGAGACAGCAGCAGCTGCTCCTGCTCTGGCAGGGTGTGGAGGTGGGAGAGCCGGGCCA
AGGCAAGGACTTCATCAGCCTCGGGCTTCAAGACGGGCACCTTGTCTTCAGGTACCAGCTGGGTGGGG
GAGGCCCGCCTGGTCTCTGAGGACCCCATCAATGACGGCGAGTGGCACCGGGTGACAGCACTGCGGGAGG
GCCGAGAGGTTCCATCCAAGTCGACGGTGAGGAGCTGGTCAGCGGCCGGTCCCCAGGTCCCAACGTGGC
AGTCAACGCCAAGGGCAGCATCTACATCGGCGGAGCCCTGACGTGGCCACGCTGACCGGGGGCAGATTTC
60 TCCTCGGGCATCACAGGCTGTGTCAAGAACCTGGTGTGCTGCACTCGGCCCGACCCGGCGCCCCGCCCCAC
AGCCCTGGACCTGCAGCACCGCGCCAGGCCGGGGCCAACACACGCCCTGCCCTCGTAGGCACCTGC

CTGCCCCACACGGACTCCCGGGCCACGCCCCAGCCCGACAATGTCGAGTATATTATTATTATTAATATTATTA
TGAATTTTTGTAAGAAACCGAGCGATGCCACGCTTTGCTCGTACCGCCCTGGGCTGGACTGGAGGTGGGC
ATGCCACCTCACACACACAGCTGGGCAAAGCCACAAGGCTGGCCAGCAAGGCAGGTTGGATGGGAGTGG
GCACCTCAGAAAGTCACCGAGACTTGGGTCAGGAACAGTGGCTGGGTGGGCCCAGAACTGCCCCCACTG
5 TCCCCCTACCCACCGATGGAGCCCCCAGATAGAGCTGGGTGGCCTGTTTCTGCAGCCCTGGGCAGTTCT
CACTCCTAGGAGAGCCAACCTCGGCTTGTGGGCTGGTGGCCCCACAGCTACCTGAGACGGGCATCGCAGGA
GTCTCTGCCACCCACTCAGGATTGGGAATTGTCTTTAGTGGCGGCTGTGGAGCAAAAGGCAGCTCACCCC
TGGGCAGGCGGTCCCCATCCCCACCAGCTCGTTTTTTCAGCACCCCCACCCACCTCCACCCAGCCCCCTGGC
CCC

10 SEQ ID NO:366

>gi|7427517|ref|NP_005520.2| heparan sulfate proteoglycan 2 (perlecan);
Heparan sulfate proteoglycan of basement membrane (perlecan) [Homo sapiens]
15 MGWRAPGALLLALLLHGRLLAVTHGLRAYDGLSLPEDIETVTASQMRWTHSYLSDDDEMLADSISGDDL
SGDLGSGDFQMVFYFRALVNFTRISIEYSPQLEDAGSREFREVSEAVVDTLESEYLLKIPGDQVVSVFIKEL
DGWVFEVLDVSGSEGNADGAQIQEMLLRVISSGSVASVYVTSPPQGFQFRRLLGTVPPFRACTEAEFACHSYN
ECVALEYRCRRPDCRDMSELNCEEPVLGISPTFSLIVETTSPLPPRPETTIMRQPPVTHAPQPLPSGSV
RPLPCGPQEAACRNHCIPRDYLCGQEDCEDGSDCLCGPPPPCEPNEFPNGHCLKLRCDGDFDC
20 EDRTDEANCPTKRPEEVCGPTQFRCVSTNMCI PASFHCDEESDCPDRSDEFGCMPPQVVTTPRESIQASR
GQTVTFTCVAIGVPAPFLINWRLNWGHIPSPRVTVTSEGGRTLIIRDVKESDQAYTCEAMNARGMVF
GIPDGVLELVQVQAGPCPDGHFYLEHSAACLPFCFGITSVCSQSTRRRFRDQIRLRFDPDDFKGVNVTMP
AQPGTTPPLSSTQLQIDPSLHEFQLVDLSRRFLVHDSFWALPEQFLGNKVDSYGGSLRYNVRYELARGMLE
PVQRPDVVLVGAGYRLLSRGHTPTQPGALNQRQVQFSEEHVHESGRPVQRAELLQVLSLEAVLIQTVY
25 NTKMASVGLSDIAMDTTTHATSHGRAHSVEECRCPIGYSGLSCESCDAHFTRVPGGPYLGTCSGCSCNG
HASSCDPVYGHCLNCQHNTTEGPQCKKAGFFGDAMKATATSCRPCPCPYIDASRRFSDTCLDLDTGQAT
CDACAPGYTGRRCESCAPGYEGNPIQPGGKCRPVNQEI VRCDERGSMGTSGEACRCKNNVGRCLNECAD
RSFHLSTRNPDGCLKCFMVGSRHCTSSWSRAQLHGASEEPGHFSLTNAASTHTTNEGI FSPTPGELGF
SSFHRLLSGPYFWSLPSRFLGDKVTSYGGELRFTVTQRSQPGSTPLHGQPLVVLQGNII ILEHHAQEPS
30 PGQPSTFIVPFREQAWQRPDGPATREHLLMALAGIDTLLIRASYAQQPAESRVSGISMDVAVPEETGQD
PALEVEQCSCPPGYRGPSQDCDGTGYTRTPSGLYLGTCECSCCHGHSEACEPETGACQGCQHHTEGPRCE
QCQPGYYGDAQRGTPQDCQLCPYGDPAAGQAHTCFLDLDGHTPCDACS PGHSGRHCERCAPGYGNPS
QQPCQQRDSQVPGPIGCNCDPQGSVSSQCDAAAGQCQCKAQVEGLTCSHCRPHHFLSASNPDGCLPCFCM
GITQQCASSAYTRHLISTHFAPGDFQGFALVNPQRNSRLTGEFTVEPVPEGAQLSFGNFAQLGHESFYWQ
35 LPETYQGDKVAAYGGKLRITLSYTAGPQGSPLSDPDVQITGNINIMLVASQPALQGPERRSYEIMFREEFW
RRPDGQPATREHLLMALADLDELLIRATFSSVPLVASISAVSLEVAQPGPSNRPRALEVEECRCPPGYIG
LSCQDCAPGYTRTGSGLYLGHCELCENGHSDLCHPETGACSQCQHNAAAGEFCELCAPGYGDATAAGTPE
DCQPCACPLTNPENMFRTCESLGAGGYRCTACEPGYTQGYCEQCGPGYVGNPSVQGGQCLPETNQAPLV
VEVHPARSIVPQGGSHSLRCQVSGRGPYFYWSREDGRPVPSGTQQRHQGSELHFPVQPSDAGVYICTC
40 RNLHRSNTSRAELLVTEAPSKPITVTVVEEQRSQSVRPGADVTFICTAKSKSPAYTLVWTRLHNGKLPTRA
MDFNGILTIRNVQLSDAGTYVCTGSNMFAMDQGTATLHVQASGTLAPVVS IHPPQLTVQPGQLAEFRCS
ATGSPTPTLEWTGGPGQLPAKAQIHGGILRLPAVEPTDQAQYL CRAHSSAGQQVARAVLHVHGGGPRV
QVSPERTQVHAGRTVRLYCRAGVPSATITWRKEGSLPPQARSERTDIATLLIPAITTADAGFYLCVAT
SPAGTAQARIQVVLASDASQPPVKIESSSPSVTEGQTLDLNLCVAGSAHAQVTWYRRGSLPHHTQVH
45 GSRLRLPQVSPADSGEYVCRVENGSGPKEASITVSVLHGTHSGPSYTPVPGSTRPIRIEPSSSHVAEGQT
LDLNCVVPQQAHAQVTHKRGSLPARHQTHGSLRLRHQVTPADSGEYVCHVVGTSGLPEASVLVTIEAS
VIPGPIPPVRIESSSSTVAEGQTLDLSCVAGQAHAQVTWYKRGGSLPARHQVGRSRLYIFQASPADAGQ
YVCRASNGMEASITVTVGTQGANLAYPAGSTQPIRIEPSSSQVAEGQTLDLNLCVVPQSHAQVTHKRG
GSLPVRHQTHGSLRLRYQASPADSGEYVCRVLGSSVPLEASVLVTIEPAGSVPALGVTPTVRIESSSQQV
50 AEGQTLDLNCLVAGQAHAQVTHKRGSLPARHQVHGSRLRLQVTPADSGEYVCRVVGSSGTQEASVLV
TIQQRLSGSHSQGVAYPVRIESSSASLANGHTLDLNLCLVASQAPHTITWYKRGGSLPSRHQIVGSRLRIP
QVTPADSGEYVCHVSNAGSRETSILVTIQSGSSHVPRVSPPIRIESSSPTVVEGQTLDLNLCVVARQPQ
AIITWYKRGGSLPSRHQTHGSHLRLHQMVSADSGEYVCRANNIDALEASIVISVSPSAGSPAPGSSMP
IRIESSSSHVAEGETLDLNCVVPQQAHAQVTHKRGSLPSYHQTRGSRLRLHHVSPADSGEYVCRVVGMS
55 SGPLEASVLVTIEASGSSAVHVPAPGGAPPRIEPPSSRVAEGQTLDLKCVVPQQAHAQVTHKRGGNLP
ARHQVHGPLRLNQNVPADSGEYSCQVTGSSGTLEASVLVTIEPSSPGPIAPAGLAQPIYIEASSSHVTE
GQTLDLNLCVVPQQAHAQVTHKRGSLPARHQTHGSLRLHHVSPADSGEYVCRAGGPGPEQEASFTVT
VPPSEGSSYRLRSPVISIDPPSSTVQQGDASFKCLIHGDAAPISLEWKTRNQELEDNVHISPNGSIITI
VGTTRPSNHGTYRCVNASNYGVAQSVVNLSPVHGPPPTVSVLPEGPVWVKVGAVTLECVSAGEPRSSARWTR
60 ISSTPAKLEQRTYGLMDSHTVLQISSAKPSDAGTYVCLAQNALGTAQKQVEVIVDTGAMAPGAPVQAE
AELTVEAGHTATLRCSATGSPARTIHWKLSPLPWQHRLEGDTLIIPRVAQQDSGQYICNATSPAGHAE
ATIILHVESPPYATTVPEHASVQAGETVQLQCLAHGTPPLTFQWSRVGSSLPGRATARNELLHFERAAPE

5

10

15

20

20

25

30

1

35

40

40

45

50

50

SEQ ID NO:371

>GL3_8D_1_M13R Direction: anti-sense

CCCCGTTGAAGTCTGATCCTTTCTTCAGTAGTTGCTGTCATGGAGTCGTCTTCCAGAATTAATGCAGCTCCTTCAG
CACCTCCAGAGCTGGAGAACTCCAACCTGTGAGTCTCATGTGTCACACACAGACCCAAAGTTCCAGGGAGCTATCA
GGTCACACAAGAAATAGCAAAGCACCTCAAAAATTTAGAAATAACCACCAAAGCCCGGGAACAAATGTGACTGCT
GCAAGAGCTTACAATCTAGGCCCTAATTCCTGTATGTGAGAGCATTTTCAAACAAAATTACTTCCCAAACAAAA
CACGTTAACAATCGAACTCATATTGAGGTCATCAACCACAGACCATAGCAGAAAACGTAGTGTTAAGTGAGACA
GAGCATAGACAGGAAAAAAGAGAAAAACAAAAGGCTCTCTGGTTTCGGCTTCTCCAGGGTGTGAGCCAATCTCC
ACCAGTTGACCAGAATCCATTGGCTGGAAAC

SEQ ID NO:372

>GL3_8D_2_M13R Direction: anti-sense

CCCCGTTGAAGTCTGGTCCTTTCTTCAGTAGTTGCTGTCATGGAGTCGTCTTCCAGAATTAATGCAGCTCCTTCAG
CACCTCCAGAGCTGGAGAACTCCAACCTGTGAGTCTCATGTGTCACACACAGACCCAAAGTTCCAGGGAGCTATCA
GGTCACACAAGAAATAGCAAAGCACCTCAAAAATTTAGAAATAACCACCAAAGCCCGGGAACAAATGTGACTGCT
GCAAGAGCTTACAATCTAGGCCCTAATTCCTGTATGTGAGAGCATTTTCAAACAAAATTACTTCCCAAACAAAA
CACGTTAACAATCGAACTCATATTGAGGTCATCAACCACAGACCATAGCAGAAAACGTAGTGTTAAGTGAGACA
GAGCATAGACAGGAAAAAAGAGAAAAACAAAAGGCTCTCTGGTTTCGGCTTCTCCAGGGTGTGAGCCAATCTCC
ACCAGTTGACCAGAATCCATTG

SEQ ID NO:373

>GL3_8D_3_M13R Direction: anti-sense

CCCCGTTGAAGTCTGGTCCTTTCTTACAGTAGTTGCTGTCATGGAGTCGTCTTCCAGAATTAATGCAGCTCCTTCA
GCACCTCCAGAGCTGGAGAACTCCAACCTGTGAGTCTCATGTGTCACACACAGACCCAAAGTTCCAGGGAGCTATC
AGGTCACACAAGAAATAGCAAAGCACCTCAAAAATTTAGAAATAACCACCAAAGCCCGGGAACAAATGTGACTGCT
TGCAAGAGCTTACAATCTAGGCCCTAATTCCTGTATGTGAGAGCATTTTCAAACAAAATTACTTCCCAAACAAAA
ACACGTTAACAATCGAACTCATATTGAGGTCATCAACCACAGACCATAGCAGAAAACGTAGTGTTAAGTGAGAC
AGAGCATAGACNGGAAAAAAGAGAAAAACAAAAGGCTCTCTGGTTTCGGCTTCTCCAGGGTGTGAGCCAATCTC
CACCAGTTGGCCAGAATCCATTGGCTGGAAACCGCGCTCAGGACGACTCGCTCTTCTCTATCCGCCGG
ACCAGCTCCACCAGCATCTCTGCCATCTTC

SEQ ID NO:374

>GL3_8D_4_M13F Direction: N/A

AATATGCTCATCCAGATCAAAACGCAGGTGGAGGCCTCGGAGGAGAGCGCCCTCAACCACCTCCAGAACCCGGGGC
GACGCGGCCGAGGGCCGGGCGGCCAAGAGGTGCGAGAAGGCCGAGGAGAAGGCCAAGGAGATTGCGAAGATGGCA
GAGATGCTGGTGGAGCTGGTCTGGCGGATAGAGAAGAGCGAGTCGTGAGCGCGGTGCGCGGTTCAGCCAA
TGGATTCTGGTCAACTGGTGGAGATTGGCTGACACCCTGGAGAAGCCGAAACCAGAGAGCCTTTTGTCTCTCTT
TTTTTCTGTCTATGCTCTGCTCACTTAACACTACGTTTTCTGCTATGGTCTGTGGTTGATGACCTCAATATG
AGTTTCGATGTGTTAACGT

SEQ ID NO:375

>GL3_8D_4_M13R Direction: anti-sense

CCCCGTTGAAGTCTGGTCCTTTCTTCAGTAGTTGCTGTCATGGAGTCGTCTTCCAGAATTAATGCAGCTCCTTCA
GGCACCTCCAGAGCTGGAGAACTCCAACCTGTGAGTCTCATGTGTCACACACAGACCCAAAGTTCCAGGGAGCTAT
CAGGTCACACAAGAAATAGCAAAGCACCTCAAAAATTTAGAAATAACCACCAAAGCCCGGGAACAAATGTGACTG
CTGCAAGAGCT

SEQ ID NO:376

>GL3_8D_PCR_G3F1 Direction: N/A

ATATGCTCATCCAGATCAAAACGCAGGTGAGGCCCTCGGAGGAGAGCGCCCTCAACACCTCCAGAACCCGGGGCGA
CGCGGCGAGGCGGGCGGGCAAAAGTGCGAGAAGGCCGAGGAGAAGGCCAGGAGATTGCGAAGATTGCAGAGAT
GCTGGTGGAGCTGGTCCGGCGGATAGAGAAGAGCGAGTCGTGAGCGCGGTGCGCGGTTCAGCCCATGGAT
TCTGGTCAACTGGTGGAGATTGGCTGACACCCTGGAGAAGCCGAAACCAGAGAGCCTTTTGTCTCTCTTTT
TCCTGTCTATGCTCTGTCTCACTTAACACTACGTTTTCTGCTATGGTCTGTGGTTGATGACCTCAATATGAGTTT
CGATTGTTAACGTGTTTTTGTGGGAAGTAATTTGTTTGAAATGCTCTCACATACAGGAATTA

SEQ ID NO:377

>gi|15193293|ref|NM_033296.1| Homo sapiens T-cell activation protein
(PGR1), mRNA

TTGTCCGTGGCTTCTCTGAGAAGAAAAGTTGAAAAGGGTAAAAGTTTTCAGGAATATTCCGGGCTCTCTA

TTGCTAAGCATAGCGAGTGTCTCGGTTTTCTCTCTCCAACAGACATCGCTATTGCGGTTCCGAGGCGAGTGGG
AAGAGATGCGGCCCCCTGGACATCGTTCGAGCTGGCGGAACCGGAGGAAGTGGAGGTGCTGGAGCCCCGAGGA
GGATTTTCGAGCAGTTTTCTGCTCCCGGTCACTCAACGAGATGCGCGAGGACATCGCGTCTGCTGACGCGCGAG
CACGGGCGGGCGTACCTGCGGAACCGGAGCAAGCTGTGGGAGATGGACAATATGCTCATCCAGATCAAAA
5 CGCAGGTGGAGGCCCTCGGAGGAGAGCGCCCTCAACCACCTCCAGAACCCGGGCGACGCGGCCGAGGGCCG
GGCGGCCAAGAGGTGCGAGAAGGCCGAGGAGAAGGCCAAGGAGATTGCGAAGATGGCAGAGATGCTGGTG
GAGCTGGTCCGGCGGATAGAGAAGAGCGAGTCTGTCGTGAGCGCGGTGCGCGGTTTCCAGCCAATGGATTTC
TGGTCAACTGGTGGAGATTGGCTGACACCCCTGGAGAAGCCGAAACCAGAGAGCCTTTTGTCTCTCTTTT
10 TTCTGTCTATGCTCTGTCTCACTTAACACTACGTTTTCTGTCTATGGTCTGTGGTTGATGACCTCAATAT
GAGTTTCGATTGTAAACGTGTTTTTGTGGGAAGTAATTTGTTTGAAAATGCTCTCACATACAGGAAT
TAGGGCCTAGATTGTAAGCTCTTGACAGTACATTTGTTCCCGGCTTTGGTGGTTATTTCTAAATTT
TTGAGGTGCTTTGCTATTTCTTGTGTGACCTGATAGCTCCCTGGAACTTTGGGTCTGTGTGTGACACATG
AGACTCACAGTTGGAGTTCTCCAGCTCTGGAGGTGCTGAAGGAGCTGCATTAATTCTGGAAGACGACTCC
15 ATGCAGCAACTACTGAAGAAAGGACCAGACTTCAACGGGGAGTGTGGATGGGCGGACCTGGCTGGGACTC
GTGAATCTGGAGAAGAGCTGGAGAATGGATAGTATTGTCTGTATTGAGACTTTAATTTCTGTGTGAGA
CCAAAGGAGGAGAGATGTGTTTTGTTCAAAATTTAAATTTGTTGTGGTACACTATCTTATGTAACCTGTC
TGGTGGATTTGTTTGGACAACCTAATCTCAGCTTTATTTGACATGGAACCTAAAATAGAAGATAAGATCTT
GATATTCTGTACAAGTTGATGTAATACCCTGATGCGTTTTAGAGGACTTGGCATAAAATGAAAGATTGGC
AAAGGCCCTTGAGGGGCTTGGGGATGACAGTATGGAACGTCTGTCATTGGACCCTAACTGGACTAGAAG
20 AGGCATCTTCAAGGTTTATACGTTGTCCAGCTGTAAGTTTCAATTTGAGTAGCAGACCTAACAAATATTTGA
GGTCAGAACCCTACCATGTTAAACAAACAAAACTTACCATGTTAATAAAAGTATTTCATTTGC

SEQ ID NO:378

>gi|15193294|ref|NP_150638.1| T-cell activation protein [Homo sapiens]
25 MRPLDIVELAEPEVEVLEPEEDFEQFLLPVINEMREDIASLTREHGRAYLRNRSKLWEMDNMLIQIKTQ
VEASEESALNHLQNPGDAAEGRAAKRCEKAEKAKEIAKMAEMLVELVRRIEKSESS

SEQ ID NO:379

Phosphoribosyl pyrophosphate synthetase-associated protein 1

>GL3_38B_PCR_G3F1 Direction: N/A

CGCCCGCCCGGTTCCCCCGCCGGCCATGAACGCCGCTCGCACCGGCTACCGAGTCTTCTCGGCCAACTCCACGGC
CGCTGTCACGGAGCTGGCCAAAGCGCATCACAGAGCGCCTTGGTGCTGAATTGGGGAAGTCTGTTGTATATGAAGA
GACCAATGGAGAAACAAGAGTTAAATAAAAAGAAATCTGTTTCGTGGCCAAGATATTTTCAATATACAGACAATACC
35 CAGAGATGTGAATACAGCTGTGATGGAGTTGCTCATCATGGCTTACGCACTGAAGACTGCCTGTGCCAGGAACAT
TATTGGGGTTCATCCCTACTTCCCTTACAGCAAGCAGAGCAAGATGAGAAAGAGGGGTTCCATTGTGTGCAAGCT
GCTAGCATCCATGCTGGCGAAAGCAGGTTTAACTCACATTATCACTATGGATCTTCATCAAAAGGAAATACAAGG
CTTTTTTCAGCTTTCCTGTGGACAACCTTAGAGCCTCACCTTTCCTGCTTCAGTATATCCGGAAGAAATCCAAAT
TACAGAAATGCAGTCATTGTAGCTAGTCTCCTGATGCTGCAAGAGGGGCCAGTCCTATGCGGAGAGACTGCGTC
40 TGGGTTTTGGCCGTCATTACCGGGGAAGCTCAGTGCACGGAACCTGG

SEQ ID NO:380

>GL3_8K_2_M13R Direction: sense

CTGCCCGCCCGGTTCCCCCTGCCNGGCCATGAACGCCGCTCTGCACCGGCTACCGAGTCTTCTCGGCCAACTC
CACGGCCGCTGCACGGAGCTGGCCAAGCGCATCACAGAGCGCCTTGGTGCTGAATTGGGGAAGTCTGTTGTATA
45 TGANAGAGACCAATGGAGGAAACCAAGAGTTACACANAGTACAAAGAANTCTGTTTCGTTGGCCAAGGATATTTTC
ATTATACAGACAATTAACCCAGAGATGTGAATACAGCTGTGATGGAGTTGCTCATCATGGCTTACGCACTGAAGA
CTGCCTGTGCCAGGAAC

SEQ ID NO:381

>GL3_8K_3_M13R Direction: sense

CGCCCGCCCGGTTCCCCCGCCGGCCATGAACGCCGCTCGCACCGGTACCGAGTCTTCTCGGCCAACTCCACGGCGCC
TGACGGAGCTGGCCAAGCGCATACAGAGCGCCTTGGTGCTGAATTGGGGAAGTCTGTTGTATATTAAGAGACCAA
TGGGAAACAAGAGTTAAATAAAAAGATCTGTTTCGTGG

SEQ ID NO:382

>GL3_8K_4_M13F Direction: anti-sense

CGCCCGCCCGNGTTTCCCCCGCCGGCCATGAACGCCGCTCGCACCCGGTACCGAGTTCTTCTCGGCCAACTCCAC
GGCCGCTGCACGGAGCCTTGGCCAAGCCGCATCACAGAGCGCCTTGGTGCTGAATGGGGAAGTCTGTGGTATA
TGAAGAGACCAATGGAGGAAACAAGAGTTAAATAAAAAGAAATTTCTGTTTCGTGGCCAAGATTATTTTTCATTATA
60 CAGACAATTACCCAGAGATGTTGAATTACAGCTTGTGATTGGAGTTTGCTCATTATGGCTTTACGCCACTGAA

GACTTGCCTTGTGTCAGGAACATTTATTGGGGTCATCCCCTTACTTCCCCTACAGCAAGCAGAGCAAAGATGAG
AAAGAGGGGTTCCATTGTGTGCAAGCTGCTTAGGCAATCCATGCTGGCCGAAAAGCNGGTTTAACTCACATTATC
ACTATGGATCTTCATCAAAGGAAATACAAGGGCTTTTACAAGCTTTCTTGTGGAACAAACCTTTTAGAAGCCT
TCAACCTTTTTCCTGCTTCAGTATATCCCAGGAAGGAAATTCCAAATTACCAGAAATGCAGGTCAATTTGGTAGCT
5 AAGTCTCCCTGATGCCTTGCCAAGGANGGGCCAGTCTATTTCGGAGAANACATGCGTCTGGTTGGGCCGGAANTC
ACACGGGAAACCNCATGCACAGAACTCNGGACCTTGG

SEQ ID NO:383

>GL3_8K_PCR_G3F1 Direction: N/A

10 CGCCCCGCCCCGGTTCCCCGCCGCCATGAACGCCGCTCGCACCGGCTACCGAGTCTTCTCGGCCAACTCCACGGC
CGCCTGCACGGAGCTGGCCAAGCGCATCACAGAGCGCCTTGGTGTGAATTGGGGAAGTCTGTTGTATATGAAGA
GACCAATGGAGAAACAAGAGTTAAATAAAAGAATCTGTTCTGGGCCAAGATATTTTCATTATACAGACAATACC
CAGAGATGTGAATACAGCTGTGATGGAGTTGCTCATCATGGCTTACGCACTGAAGACTGCCTGTGCCAGGAACAT
TATTGGGGTCATCCCCTACTTCCCCTACAGCAAGCAGAGCAAGATGAGAAAGAGGGGTTCCATTGTGTGCAAGCT
15 GCTAGCATCCATGCTGGCGAAAGCAGGTTTAACTCACATTATCACTATGATCTTCATCAAAAGGAAATACAAGGC
TTTTTTCAGCTTTCTGTGGACAACCTTAGAGCCTCACCTTTCCTGCTTCAGTATATCCGGAAGAAATTCCAAATT
ACAGAAATGCAGTCATTGTAGCTAACGTCTCCTGATGCTGCAAGAGGGGCCAGTCCCTATGCGGAGAGACTGCGT
CTGGGTTTGGCCGTCATTACGGGGAAGCTCAGTGCACGGAAGTGGACATTGGACGATGGTTCGTCAACTCCCCGC
20 TATGGTCAAAATGCTACTGTGCACCC

SEQ ID NO:384

>gi|4506130|ref|NM_002766.1| Homo sapiens phosphoribosyl pyrophosphate
synthetase-associated protein 1 (PRPSAP1), mRNA

25 GGTGCGCAAGGGCACGGACCTCGGAGCTCTCCCCGTTCCCCCGCCGCCATGAACGCCGCTCGCACCGGC
TACCGAGTCTTCTCGCCAACCTCCACGGCCGCTGCACGGAGCTGGCCAAGCGCATCACAGAGCGCCTTG
GTGCTGAATTGGGGAAGTCTGTTGTATATCAAGAGACCAATGGAGAAACAAGAGTTGAAATAAAAGAATT
TGTTCTGGCCAAGATATTTTCATTATACAGACAATCCAGAGATGTGAATACAGCTGTGATGGAGATTG
CTCATCATGGCTTACGCACTGAAGACTGCCTGTGCCAGGAACATTATTGGGGTCATCCCCTACTTCCCCT
ACAGCAAGCAGAGCAAGATGAGGAAGAGGGGTTCCATTGTGTGCAAGCTGCTAGCATCCATGCTGGCGAA
30 AGCAGGTTTAACTCACATTATCACTATGGATCTTCATCAAAAGGAAATACAAGGCTTTTTTCAGCTTTCT
GTGGACAACCTTAGAGCCTCACCTTTCTGCTTCAGTATATCCAGGAAGAAATTCCAAATTACAGAAATG
CAGTCATTGTAGCTAAGTCTCCTGATGCTGCAAGAGGGGCCAGTCCCTATGCGGAGAGACTGCGTCTGGG
TTTGGCCGTCATTACGGGGAAGCTCAGTGCACGGAAGTGGACATGGACGATGGTTCGTCCTCCCCGCT
ATGGTCAAAAATGCTACTGTGCACCCAGGCTGGAGTTGCCATTGATGATGGCCAAAGAGAAGCCACCGA
35 TAACTGTAGTTGGAGATGTTGGAGGCGCATCGCAATCATCGTGGATGACATTATTGACGATGTGGAGAG
TTTTGTTGCTGCCGCGGAGATCCTGAAAGAGAGAGGCGCCTATAAGATCTATGTTATGGCCACCCACGGC
ATCCTGTCTGCAGAGGCCCTCGCCTGATTGAGGAGTCTCCGTAGACGAGGTGGTGGTGACGAATACATG
TCCCTCATGAGGTTTCAAGAGCTGCAATGTCCCAAGATAAAGACTGTGGATATCAGTTTGATTCTTTCTGA
AGCCATTCCGAGAATCCACAATGGAGAGTCCATGGCCTACCTTTTCCGAAACATCACTGTGGATGACTAG
40 CTTTTCACGAGGGTCTCGACCTGGACCTCCTGAGGGAACATGGAAAAGCAGTGCCATGAGTGATACAG
TGTTTCTTGAAGGGAGGACTCGAAACAGCCTGGAGTTAGATATCTTCTTTTGCCCGGATTGATGGGGA
GGAGGGATTAAAAGAGTCAGGAAGAAGACAGAGCTAATGGATAAATATCATAACATGGCCTTACATGTCT
GCTGTCTACAGCCCTGTTCTTAAAGTTCTAGCTGCTTTCTTAAATAATCTGAAAATCTTATTGATA
CTAAAGAGGAGTTAAAGGCACATAAAGTCTTAACTCTATAATGTTTCAATTTAGTTGTTTCAGCTCCAGGGA
45 AATGGAGGTATTGATGTTGAACCTGGTTAGGGAAGCTGAGCGCCTGTGGCCCTATTACTATCCAGTTGGC
CTCTCCCAAATCAACTTCAAGTCTTTTATAGAGAATCGTATTTTCTTTCAGAAATTGCTATGCCTACAG
CCATTGAAAATGAAGCATTCATGTTTACATCTTCCAAGGATGTCAGATTAGAAAATAGCATCCCACC
TCTGGGTATCTGAGTGGCTCTGAAGTTGCAAATAAAATAATTTGTTGT

SEQ ID NO:385

>gi|4506131|ref|NP_002757.1| phosphoribosyl pyrophosphate synthetase-
associated protein 1 [Homo sapiens]

55 MNAARTGYRVFLANSTAACTELAKRITERLGAEIGKSVVYQETNGETRVEIKEFVRGQDIFIIQTIPRDV
NTAVMELLIMAYALKTACARNIIGVIPYFPYSKQSKMRKGSIVCKLLASMLAKAGLTHIITMDLHQKEI
QGFFSFPVDNLRASPFLLQYIQEIPNYRNAVIVAKSPDAAKRAQSYAERLRLGLAVIHGEAQCTELDMD
DGRHSPPMVKNAIVHPGLELPLMMAKEKPPITVVGDVGGRIAIIVDDIIDDVESFVAAAEILKERGAYKI
YVMATHGILSAEAPRLIESSVDEVVVTNTVPHEVQKLQCPKIKTVDISLILSEAIRRIHNGESMAYLFR
NITVDD

SEQ ID NO:386

Plakoglobin

>G3_1_90_PCR_G3F1 Direction: N/A

GTGCGTACCATGCAGAATACCAGCGACCTGGACACAGCCCGCTGCACCACCAGCATCCTGCACAACCTCTCCAC
5 CACCGGGAGGGGCTGCTCGCCATCTTCAAGTCGGGTGGCATCCCTTCTCTGGTCCGCATGCTCAGCTCCCCCTTT
GAGTCGGTCTCTTCTATGCCATCACCACGCTGCACAACCTGCTCCTGTACCAGGAGGGCGCCAAGATGGCCGTG
CGCCTGGCCGACGGGCTGCAAAAGATGGTGCCCTGCTCAACAAGAACAACCCCAAGTTCTGGCCATCACCACC
GACTGCCTGCAGCTCCTGGCCTACGGCAACCAGGAGAGCAAGCTGATCATCCTGGCCAATGGTGGGCCCCAGGGC
10 CCTCGGTGCACGATTATGCGTTACTTCTGTTTAAAAGC

SEQ ID NO:387

>G3_4_01_PCR_G3F1 Direction: N/A

CCAGTGCTGTGCTTGCCTACCATGCAGAATACCAGCGACCTGGACACAGCCCGCTGCACCACCAGCATCCTGCAC
AACCTCTCCACACCGGGAGGGGCTGCTCGCCATCTTCAAGTCGGGTGGCATCCCTGCTCTGGTCCGCATGCTC
15 AGTCCCCCTTTTGAGTCGGTCTCTGTTCTATGCCATCACCACGCTGCACAACCTGCTCCTGTACCAGGAGGGCGC
CAAGATGGCCGTGCGCCTGGCCGACGGGCTGCAAAAGATGGTGCCCTGCTCAACAAGAACAACCCCAAGTTCTG
GCCATCACCACCGACTGCCTGCAGCTCCTGGCCTACGGCAACCAGGAGAGCAAGCTGATCATCCTGGCCAATGGT
GGGCCCCAGGCCCTCGTGCAGATCATGCGTAACTACGTTATTTAAAAGCTGCTC

SEQ ID NO:388

>G3_4_44_PCR_G3F1 Direction: N/A

TCCAGTGCTGTGCTGCGTACCATGCAGAATACCAGCGACCTGGACACAGCCCGCTGCACCACCAGCATCCTGCAC
AACCTCTCCACACCGGGAGGGGCTGCTCGCCATCTTCAAGTCGGGTGGCATCCCTGCTCTGGTCCGCATGCTC
AGTCCCCCTTTTGAGTCTGTCTCTGTTCTATGCCATCACCACGCTGCACAACCTGCTCCTGTACCAGGAGGGCGC
25 CAAGATGGCCGTGCGCCTGGCCGACGGGCTGCAAAAGATGGTGCCCTGCTCAACAAGAACAACCCCAAGTTCTC
GGCCATCACCACCGACTGCCTGCAGCTCCTGGCCTACGGCAACCAGGAGAGCAAGCTGATCATCCTGGCCAATGG
TGGGCCCCAGGCCCTCGTGCAGATCATGCGTAACTACGTTATGAAAAGCTGCTCTGGACCACCAGTCGTGTGCTC
AAGGTGCTATCCGTGTGTCCAGCAATAAGCCTGCCATTGTGGAGGCTGGTGGGATGCAGGCCCTGGGCAAGCAC
CTGACCAGCAACAGCCCCCGCCTGGTGCAGAACTGCCTGTGGACCCTGCGCAACCTCTCAGATTTGGCCACCAGG
30 CAGGAGGGC

SEQ ID NO:389

>G3_4_90_PCR_G3F1 Direction: N/A

CCAGTGCTGTGCTGCGTACCATGCAGAATACCAGCGACCTGGACACAGCCCGCTGCACCACCAGCATCCTGCACA
ACCTCTCCACACCGGGAGGGGCTGCTCGCCATCTTCAAGTCGGGTGGCATCCCTGCTCTGGTCCGCATGCTCA
35 GCTCCCCCTTTGGGAGTCGGTCTCTGTTCTATGCCATCACCACGCTGCACAACCTGCTCCTGTACCAGGAGGGCGCC
AAGATGGCCGTGCGCCTGGCCGACGGGCTGCAAAAGATGGTGCCCTGCTCAACAAGAACAACCCCAAGTTCTC
GCCATCACCACCGACTGCCTGCAGCTCCTGGCCTACGGCAACCAGGAGAGCAAGCTGATCATCCTGGCCAATGGT
GGGCCCCAGGCCCTCGTGCAGATCATGCGTAACTACGTTATGAAAAGCTGCTCTGGACCACCAGTCGTGTGCTC
40 AAGGTGCTATCCGTGTGTCCAGCAATAAGCCTGCCATTGTGGAGGCTGGTGGGATGCAGGCCCTGGGCAAGCAC
CTGACCAGCAACAGCCCCCGCCTGGTGCAGAACTGCCTGTGGACCCTGCGCAACCTCTCAGATTTGGCCACCAGG

SEQ ID NO:390

>G3_5_91_PCR_G3F1 Direction: N/A

TGCTGTGCTGCGTACCATGCAGAATACCAGCGACCTGGACACAGCCCGCTGCACCACCAGCATCCTGCACAACCT
CTCCACACACCGGGAGGGGCTGCTCGCCATCTTCAAGTCGGGTGGCATCCCTGCTCTGGTCCGCATGCTCAGCTC
CCCTTGGAGTCGGTCTCTGTTCTATGCCATCACCACGCTGCACAACCTGCTCCTGTACCAGGAGGGCGCCAAGATG
GCCGTGCGCCTGGCCGACGGGCTGCAAAAGATGGTGCCCTGCTCAACAAGAACAACCCCAAGTTCTGGCCATC
ACCACCGACTGCCTGCAGCTCCTGGCCTACGGCAACCAGGAGAGCAAGCTGATCATCCTGGCCAATGGTGGGCCC
50 CAGGCCCTCGTGCAGATCATGCGTAACTACAGTTATGAAAAGCTGCTCTGGACCACCAGTCGTGTGCTCAAGGTG
CTATCCGTGTGTCCAGCAATAAGCCTGCCATTGTGGAGGCTGGTGGGATGCAGGCCCTGGGCAAGCACCTACCA
GCAACAGCCCCCGCCTGGTGCAGAACTGCCTGTGGACCCTGCGCAACCTCTCAGATTTGGCCACCAGGAGGAGG
GCCTGGAGAGTGTGCTGAAAGATTCTGGTGA

SEQ ID NO:391

>G3_7_64_PCR_G3F1 Direction: N/A

AGCTTTTTATAACTGTAGTTACGCATGATCTGCACGAGGGCCTGGGGCCACCATTTGGCCAGGATGATCAGCTTG
CTCTCCTGGTTGCCGTAGCCAGAGCTGCAGCAGTCGTGTATGCCAGAACCTTGGGGTTGTTCTTGTGAGCAGGGG
CACCATCTTTTGCAGCCCGTCGCCAGGCGCACGGCCATCTTGGCGCCCTCCTGGTACAGGAGCAGGTTGTGCAGC
60 GTGGTGTATGATAGAACAGGACCGACTCCACANGGGAGCTGAGCATGCGGACCAGAGCAGGATGCCACCCACTTG

AAATGCAGCAGCCCCCTCCCGGTGTGGAGAGGTTGTGCAGGATGCTGTGGTGCAGCGGGCTGTGTCCAGTCGCTGT
ATTCTGCATGGTACGCACGACGCGCCACCACTGGGCGAGCCCATCAGGGCCCCGCGCGACCCCTCCTTTTCGACGC
TGTTTACAATCATGCCCCCTTGTAC

5 SEQ ID NO:392

>GL1_34_HIGH_1_G3F1 Direction: N/A

10 GAGAGGCGCCATNGNCGGTGGCAGCGGTGGCTCCAGTGCTGTCTGCTGCGTACCATGCAGAATACCAGCGACCTGGA
CACAGCCCGCTGCACCACCAGCATCCTGCACAACCTCTCCACCACCGGGAGGGGCTGCTCGCCATCTTCAAGTC
GGGTGGCATCCCTGNTCTTGGTCCGCATGCTCAGCTCCCTGNTGGAGTCGGTCCTGTTCTATGCCATCACCACG
15 CTGCACAACCTGCTCCTGTACCAGGAGGGCGCCAAGATGGCCGTGCGCCTGGCCGACGGGCTGCAAAAGATGGTG
CCCCTGCTCAACAAGAACAACCCCAAGTTCCCTGGCCATCACCACCGACTGCTGTCAGCTCCTGGCCTACGGCAAC
CAGGAGAGCAAGCTGATCATCCTGGCCAATGGTGGGCCCCAGGCCCTCGTGCAGATCATGCGTAACCTACAGTTAT
GAAAAGCTGCTCTGGACCACAGTCGTGTGCTCAAGGTGCTATCCGTGTGTCCCAGCAATAAGCCTGCCATTGTG
15 GAGGCTGGTGGGATGCAGGCCCTGGGCAAGCACCTGACCAGCAACAGCCCCCGCTGGTGCAGAACTGCCTGTGG
ACCTGCGCAACCTCTCAGATGTGGCCACCAGGCAGGAGGGCCTGGAG

SEQ ID NO:393

>GL1_34_HIGH_2_G3F1 Direction: N/A

20 GNCGGTGGCAGCGGTGGCTCCAGTGCTGTCTGCTGCGTACCATGCAGAATACCAGCGACCTGGACACAGCCCGCTGC
ACCACCAGCATCCTGCACAACCTCCCCACCACCGGGAGGGGCTGCTCGCCATCTTCAAGTCGGGTGGCATCCCT
GCTCTGGTCCGCATGCTCAGCTCCCTGNTGGAGTCGGTCCTGTTCTATGCCATCACCACGCTGCACAACCTGCT
CCTGTACCAGGAGGGCGCCAAGATGGCCGTGCGCCTGGCCGACGGGCTGCAAAAGATGGTGGCCCTGCTCAACAA
GAACAACCCCAAGTTCTGGCCATCACCACCGACTGCTGTCAGCTCCTGGCCTACGGCAACCAGGAGAGCAAGCT
25 GATCATCCTGGCCAATGGTGGGCCCCAGGCCCTCGTGCAGATCATGCGTAACCTACAGTTATGAAAAGCTGCTCTG
GACCACCAGTCGTGTGCTCAAGGTGCTATCCGTGTGTCCCAGCAATAAGCCTGCCATTGTGGAGGCTGGTGGGAT
GCAGGCCCTGGGCAAGCAGGAGGGCCTGGAGAGTGTGCTGAAGATTCTGGTGAATCAGCTGAGTGTGGATGA
CTCAGATGTGGCCACCAAGCAGGAGGGCCTGGAGAGTGTGCTGAAGATTCTGGTGAATCAGCTGAGTGTGGATGA
CGTCAACGTCTCCTACCTGTGCCACGGGCACACTCTCCAACCTGACATGCAACAACAGCAAGAACAAGACGCTGGT
30 GACACAGAACAGCGGTGTGGAGGCTCTCATCCATGCCATCTGCGTGTGGTGAACAAGGACGACATCACGGAGCC
TGCCGTCTGCGCTCTGCGCCACCTCACTAGCCGNCACCCCTGAGGNCGAGAACGGCCTGGCCCCAGTAACGGGCC
CTCTNTGCAGGCTTTCTCTCTCTAGAACTCCTTCTGTTGGAAGGCCCTCCG

SEQ ID NO:394

>GL1_34_HIGH_3_G3F1 Direction: N/A

35 GAGAGGCGCCATNGNCGGTGGCAGCGGTGGCTCCAGTGCTGTCTGCTGCGTACCATGCAGAATACCAGCGACCTGGA
CACAGCCCGCTGCACCACCAGCATCCTGCACAACCTCTCCACCACCGGGAGGGGCTGCTCGCCATCTTCAAGTC
GGGTGGCATCCCTGCTCTGGTCCGCATGCTCAGCTCCCTGATGGAGTCGGTCCTGTTCTATGCCATCACCACGC
TGACAACCTGCTCCTGTACCAGGAGGGCGCCAAGATGGCCGTGCGCCTGGCCGACGGGCTGCAAAAGATGGTGC
40 CCCTGCTCAACAAGAACAACCCCAAGTTCTGGCCATCACCACCGACTGCTGTCAGCTCCTGGCCTACGGCAACC
AGGAGAGCAAGCTGATCATCCTGGCCAATGGTGGGCCCCAGGCCCTCGTGCAGATCATGCGTAACCTACAGTTATG
AAAAGCTGCTCTGGACCACAGTCGTGTGCTCAAGGTGCTATCCGTGTGTCCCAGCAATAAGCCTGCCATTGTGG
AGGCTGGTGGGATGCAGGCCCTGGGCAAGCACCTGACCAGCAACAGCCCCCGCTGGTGCAGAACTGCCTGTGGA
CCCTGCGCAACCTCTCAGATGTGGCCACCAAGCAGGAGGGCCTGGAGAGTGTGCTGAAGATTCTGGTGAATCAGC
TGAGTGTGGATGACGTCAACGTCCTCACCTGTGCCACGGGCACACTCTCCAACCTGACATGCAACAACAGCAAGA
45 ACAAGACGCTGGTGACACAGAACAGCGGTGTGGAGGCTCTCATCCATGCCATCCTGCGTGTGGTGACAAGGACG
ACATCACGGAGCCTGCGCTCTGCGCTCTGNGCCANCTCATTAGCCGCCACCCCTGAGGGCGAGATGGNCCAGAACT
CTGTGCGTCNCAACTATGGCATCCCA

SEQ ID NO:395

>GL3_1.36_A3_M13F Direction: N/A

50 TGGTGGTCCAGAGCAGCTTTTTATAACTGTAGTTACGCATGATCTGCACGAGGGCCTGGGGCCCACCATTGGCCA
GGATGATCAGCTTGCTCTCCTGGTTGCCGTAGGCCAGGAGCTGNAGGCAGTCGGTGGTGTATGGCCANGAAGTTGG
GGTTGTNCTTGTGTGAGCAGGGGCACCATTTTTGTCAGCCCGTCGGCCAGGCGCACGGCCATCTTGGCGCCCTCCT
GGTACAGGAGCAGGTTGTGTCAGCGTGGTGTATGGCATAGAACAGGACCGACTCCACAGGGGAGCTGATCATGCGGA
55 CCAGAGCAGGGATGCCACCCGACTTGAAGATGGCGAGCAGCCCCCTCCCGGTGGTGGGAGAGGTTGTGTCAGGATGC
TGGTGGTGCAGCGGGCTGTGTCCAGGTCGCTGGTATTCTGCATGGTACGCACGACAGCGGCCACCAGCTGGGGCG
AGCCCATCANGGCCGNCGCGACGCCTCCTTGTTTCGACAGCTGGTCCACAATCATGGCCGNCTTGGTCAACACCA
CCGGGTCTCTCG

>GL3_1.36_A3_M13R Direction: anti-sense

5

>GL3 10.28 PCR G3F1 Direction: N/A

20

>GL3_9.60_PCR_G3F1 Direction: N/A

30

>GL3 9.66 PCR G3F1 Direction: N/A

45

>GL3 9.94 PCR G3F1 Direction: N/A

60

GGTTGTTCTTGTGAGCAGGGGCACCATCTTTTGCAGCCCGTCGGCCAGGCGCACGGCCATCTTGGCGCCCTCCT
GGTACAGGAGCAGGTTGTGTCAGCGTGGTGTATGGCATAGAACAGGACCGACTCCACAGGGGAGCTGAGCATGCGGA
CCAGAGCAGGGATGCCACCCGACTTGAAGATGGCGAGCAGCCCCCTCCCGGTGGTGGGAGAGGTTGTGTCAGGATGC
TGGTGGTGCAGCGGGCTGTGTCCAGGTCGCTGGTATTCTGCATGGTACGCACGACAGCGGCCACCAGCTGGGGCG
5 AGCCCATCAGGGGCCGCGCGCAGCCCTCCTTGTTCGACAGCTGGTTTCAATCATGGCCGCTTTGGTCACCAC
CACCGGGTCTCGTCGTTGAGCAGTTTGGTTAAGCTCGGGCAGGGCGCGAGTGGCCAGCTCGGCATCGTCCTGGT
AGTTGATGAGATGCACAA

SEQ ID NO:401

>GL1_34_HIGHPLKS_1 Direction: N/A

CANGAGAGCAAGCTGATCATCTTGGCCAATGGTGGGCCCCAGGCCCTCGTGCAGATCATGCGTAACTACAGTTAT
GAAAAGCTGCTCTGGACCACCAGTCGTGTGCTCAAGGTGCTATCCGTGTGTCCAGCAATAAGCCTGCCATTGTG
GAGGCTGGTGGGATGCAGGCCCTGGGCAAGCACCTGACCAGCAACAGCCCCGACCTGGTGCAGAACTGCCTGTG
GACCCCTGCGCAACCTCTCAGATGTGGCCACCAGGCAGGAGGGCCTGGAGAGTGTGCTGAAGATTCTGGTGAATCA
15 GCTGAGTGTGGATGACGTCAACGTCTCCTGTCACCGGCCACACTCTCCAACCTGACATGCAACAACAGCAA
GAACAAGACGCTGGTGACACAGAACAGCGGTGTGGAGGCTCTCATCCATGCCATCCTGCGTGTGGTGACAAGGA
CGACATCACGGAGCCTGCCGTCTGCGCTCTGCGCCACCTCACTAGCCGCCACCCTGAGGCCGAGATGGCCAGAA
CTCTGTGCGTCTCAACTATGGCATCCAGCCATCGTGAAGCTGCTCAACCAGCCCCAACCAGTGGCCACTGGTCAA
GGCAACCATCGGCTTGATCAGGAATCTGGCCCTGTGCCCCAGCCCAACCATGCCCGCTGCAGGAGGCAGNGGTC
20 ATCCCCGCTCGTCCAACCTGCTGGTGAANGCCACCAGGATGCCCAGNGCCACGTANCTGCAGGCNCACCGCAGC

SEQ ID NO:402

>GL1_34_HIGHPLKS_2 Direction: N/A

CCCTGNATAATACNTTAGTGTGGATGACGTCAANGTCCTCINNCTGCTGCCACGGGGACACTCTCCAACCTGACAT
25 GCAACAACAGCAAGAACAAGACGCTGGTGACACAGAACAGCGGTGTGGAGGCTCTCATCCATGCCATCCTGCGAT
GCTGGTGACAAGGACGACATCACGGAGCCTGCCGTCTGCGCTCTGCGCCACCTCACTAGCCGCCACCCTGAGGCC
GAGATGGCCAGAACTCTGTGCGTCTCAACTATGGCATCCAGCCATCGTGAAGCTGCTCAACCAGCCCCAACCAG
TGGCCACTGGTCAAGGCAACCATCGGCTTGATCAGGAATCTGGCCCTGTGCCCAGCCCAACCAGTGGCCAG
GAGGCAGCGGTATCCCCCGCTCGTCCAACCTGCTGGTGAAGGCCACCAGGATGCCAGCGCCACGTAGCTGCA
30 GGCACACAGCAGCCCTACACGGATGGTGTGAGGATGGAGGAGATTGTGGAGGGCTGCACCGGAGCACTGCACATC
CTCGCCCGGGACCCCATGAACCGCATGGAGATCTTCCGGCTCAACACCATTCCCCTGTTTGTGTCAGCTCCTGTAC
TCGTGGTGGAGAACATCCAGCGGTGGCTGCCGGGGTGTGTGTGAGCTGGCCAGGACAAGGAGGCGGTGCGAC
GCCATTGATGCGAGGGGGCCTCGGCCCACTCATGGAGTTGCTGCACTCCCGCAACGAGGGCACTGCCACCTAC
GCTGCTGCCCGTCTGTTCCGCAATCTCCGANGACAANAACCCAGACTACCGGAANCGCGTGTCCGTGGAGCTCAC
35 CAACTCCCTTCTTCANGCATGAACCCGGCTGCC

SEQ ID NO:403

>GL1_34_HIGHPLKS_3 Direction: N/A

GCACACTGNACTGATCNGGTATCTGGCCCNNTGCCAGCCAACCATGCCCGCTGCAGGAGGCAGCGGATCATCC
40 CCCGCTCGCTCCAACCTGCTGGTGAAGGCCACCAGGATGCCAGCGCCACGCTAGCTGCAGGCACACAGCAGCC
CTACACGGATGGTGTGAGGATGGAGGAGATTGTGGAGGGCTGCACCGGAGCACTGCACATCCTCGCCCGGGACCC
CATGAACCGCATGGAGATCTTCCGGCTCAACACCATTCCCCTGNTTTGTGTCAGCTCCTGTACTCGTCGGTGGAGA
ACATCCAGCGCTGGCTGCCGGGGTGTGTGTGAGCTGGCCAGGACAAGGAGGCGGCCGACGCCATTGATGCAG
AGGGGGCCTCGGCGCCACTCATGGAGTTGCTGCACTCCCGCAACGAGGGCACTGCCACCTACGCTGCTGCCGTCC
45 TGTTCCGCATCTCCGAGGACAAGAACCAGACTACCGGAAGCGCGTGTCCGTGGAGCTCACCAACTCCCTCTTCA
AGCATGACCCGGCTGCCCTGGGAGGCTGCCAGAGCATGATTCCCATCAATGAGCCCTATGGAGATGACTNGGATG
CCACCTACCGCCCCATGTACTCCAGCGATGTGCCCTTGAACCCGCTGGAGATGCACATGGACATGGATGGAGAC
TACCCCATCGACACCTACAGNGANGGCCCTCAGGCCCGGTACCCCACTGNAGACCACATGCTGNNCTACGCGGCC
TGGCCCCAGTACGNNCCCTCTTTGAGGCTTTTCTCTCTCTAGAAACCTCCTTCTGTTGGAGGCCCTCNCAA
50 AACCCANCAGNACAACCCACCACAGNGGTTACATAGATAAAGCGNCCGNTCGACTANTCTGAGGTCTGATACTCA
CTGACTGTCNTACACAGTGAACCTATAAANTAAAA

SEQ ID NO:404

>GL1_34_HIGHPLKS_4 Direction: sense

AACCTNTGAATGNCGACCCCTTGTGTGAGCTGGCCCAAGGACAAGGCAGGCGGCCGACGCCATTGATGCANAGGG
55 GGCCTCGGCCCCACTCATGGAGTTGCTGCACTCCCGCAACGAGGGCACTGCCACCTACGCTGCTGCCGTCTGTT
CCGCATCTCCGAGGACAAGAACCAGACTACCGGAAGCGCGTGTCCGTGGAGCTCACCAACTCCCTCTTCAAGCA
TGACCCGGCTGCCCTGGGAGGCTGCCAGAGCATGATTCCCATCAATGAGCCCTATGGAGATGACTTGGATGCCAC
CTACCGCCCCATGTACTCCAGCGATGTGCCCTTGACCCGCTGGAGATGCACATGGACATGGATGGAGACTACCC

CATCGACACCTACAGCGACGGCCTCAGGCCCCCGTACCCCACTGCAGACCACATGCTGGCCTAGGCGGCCTGGCC
CCAGTACGGCCCCCTCTTTGCAGGCTTTTCTCTCTCTAGAACCTCCTTCTGTTGGAGGCCCT

SEQ ID NO:405

>GL1_34_HIGH_1_G3R1 Direction: N/A

AGGGCCTCCAACAGAAGGAGGTTCTAGAGAGGAGGAAAAGCCTGCAAAGAGGGGGCCGTACTGGGGCCAGGCCGC
CTAGGCCANCATGTGGTCTGCAGTGGGGTACGGGGGCCCTGAGGCCGTGCTGTAGGTGTGATGGGGTAGTCTCC
ATCCATGTCCATGTGCATCTCCAGNNGGTCAAGGGGCACATCGCTGGAGTACATGGGGCGGTAGGTGGCATCCAA
GTCATCTCCATAGGGCTCATTGATGGGAATC

SEQ ID NO:406

>gi|4504810|ref|NM_002230.1| Homo sapiens junction plakoglobin (JUP),
transcript variant 1, mRNA

CGCCAGAGTCCGGAGCAGCCGCGCCGACCGAGCTCAGTTCGCTGTCCGCGCCGGCTCCCACCC
CGGCCCGACCCCGACCCGGCCCGGTACAGGCCCATACTCAGTAGCCACGATGGAGGTGATGAACCTGATG
GAGCAGCCTATCAAGGTGACTGAGTGGCAGCAGACATACCTACGACTCGGGTATCCACTCGGGCGCCA
ACACCTGCGTGCCTCCGTCCGTGAGCAGCAAGGGCATCATGGAGGAGGATGAGGCCTGCGGGCGCCAGTACAC
GCTCAAGAAAACCACCACTTACACCCAGGGGGTGCCTCCAGCCAAGGTGACCTGGAGTACCAGATGTCC
ACAACAGCCAGGGCCAAACGGGTGCGGGAGGCCATGTGCCCTGGTGTGTGTCAGGCGAGGACAGCTCGCTTC
TGCTGGCCACCCAGGTGGAGGGGACGGCCACCAACCTGCAGCGACTGGCCGAGCCGTCCCAGCTGCTCAA
GTCGGCCATTGTGCATCTCATCAACTACCAGGACGATGCCGAGCTGGCCACTCGCGCCCTGCCCCGAGCTC
ACCAAACCTGCTCAACGACGAGGACCCGGTGGTGGTGACCAAGGCGGCCATGATTGTGAACCAAGCTGTGCA
AGAAGGAGGCGTTCGCGCGGGGCCCTGATGGGCTCGCCCCAGCTGGTGGCCGCTGTGCTGCGTACCATGCA
GAATACCAGCGACCTGGACACAGCCCGCTGCACCAACAGCATCCTGCACAACCTCTCCACCCAGGGAG
GGGTGCTCGCCATCTTCAAGTCGGGTGGCATCCCTGCTCTGGTCCGCATGCTCAGCTCCCCCTGTGGAGT
GGTCTCTGTTCTATGGCATCACACGCTGCACAACCTGCTCCTGTACCAGGAGGGCGCCAAGATGGCCGT
GCGCCTGGCCGACGGGCTGCAAAAGATGGTGGCCCTGCTCAACAAGAACAACCCCAAGTTCTTGGCCATC
ACCACCGACTGCCTGCAGCTCCTGGCCTACGGCAACCAGGAGAGCAAGCTGATCATCCTGGCCAATGGTG
GGCCCCAGGCCCTCGTGCAGATCATGCGTAACCTACAGTTATGAAAAGCTGCTCTGGACCACCAAGTCTGT
GCTCAAGGTGCTATCCGTGTGTCCAGCAATAAGCCTGCCATTGTGGAGGCTGGTGGGATGCAGGCCCTG
GGCAAGCACCTGACCAGCAACAGCCCCCGCCTGGTGCAGAACTGCCTGTGGACCCCTGCGCAACCTCTCAG
ATGTGGCCACCAAGCAGGAGGGCCTGGAGAGTGTGCTGAAGATTCTGGTGAATCAGCTGAGTGTGGATGA
CGTCAACGTCCTCACCTGTGCCACGGGCACACTCACAACCTGACATGCAACAACAGCAAGAACAAGACG
CTGGTGACACAGAACAGCGGTGTGGAGGCTCTCATCCATGCCATCCTGCGTGTGGTGAACAAGGACGACA
TCACGGAGCCTGCGCTGTGCGCTCTGCGCCACCTCACTAGCCGCCACCCCTGAGGCGGAGATGGCCAGAA
CTCTGTGCGTCTCAACTATGGCATCCAGCCATCGTGAAGCTGCTCAACCAGCCCAACCAAGTGGCCACTG
GTCAAGGCAACCATCGGCTTGATCAGGAATCTGGCCCTGTGCCCAGCCAACCATGCCCCGCTGCAGGAGG
CAGCGGTCTATCCCCCGCCTCGTCCAACCTGCTGGTGAAGGCCACCAAGGATGCCAGCGCCACGTAGCTGC
AGGCACACAGCAGCCCTACACGGATGGTGTGAGGATGGAGGAGATTGTGGAGGGCTGCACCGGAGCACTG
CACATCCTCGCCCGGGACCCCATGAACCGCATGGAGATCTTCCGGCTCAACACCATTCCCCTGTTTGTG
AGCTCCTGTACTCGTGGTGAGAACATCCAGCGGTGGCTGCCGGGGTGTGCTGTGTGAGCTGGCCAGGA
CAAGGAGGCGGCCGACGCCATTGATGCAGAGGGGGCTCGGCCCACTCATGGAGTTGCTGCACTCCCGC
AACGAGGGCACTGCCACCTACGCTGCTGCCCTCTGTTTCCGCTCTCCGAGGACAAGAACCAGACTACC
GGAAGCGCGTGTCCGTGGAGCTCACCAACTCCCTCTTCAAGCATGACCCGGCTGCCTGGGAGGCTGCCCA
GAGCATGATTCCCATCAATGAGCCCTATGGAGATGACATGGATGCCACCTACCGCCCCATGTACTCCAGC
GATGTGCCCCCTTGACCCGCTGGAGATGCACATGGACATGGATGGAGACTACCCCATCGACACCTACAGCG
ACGGCCTCAGGCCCCCGTACCCCACTGCAGACCACATGCTGGCCTAGGCGGCCTGGCCCCAGTGACGGCC
CCCTCTTTGACAGGCTTTTCTCTCTCTAGAACCTCCTTCTGTTGGAGGCCCTCCCATCTCCCCGCTGAA
ACCTGCGCTCCTTTTTTGGGGGGATCCTTTGCTGCTGAGCTTCCCCAAGCACGGTGTGCCCTGGCCTGCC
TTCTTCTTGTGTCTTTGGTGGGGATGGGGAGGCCATTCTCTGCTGGCCCTTCTGGGGGTGGTGGGCAGG
TGACACGGAGTGGCTTGAGCTTCTGGGGATGCAGGTCCACCGAGCCCTGACCCCTGTCTGTCCCCGCTC
CCCTAACAGGTGCGGTTCCTCATCTGAGAGGCTCTCCGTGCAGGCGATGGGGCAAGACAGAAAAGTGCTT
GAGCTGGGGAAAGCCGGGGTGTAACTTCTGCTGCACCTGCGCCTCCAGAGGTCTCCGTAGGGTCTTTT
TTGGGATAGTGTCTGCTCCTGCTTTTCTGCTCTGGGCATGGGTCCAGGGCCTGACACCCCTCCCCGCC
CCTGTGGCCCTGGCCACTAAAGCTTCAGACTCAAGTACCCATTCTGTTTTCCCCCAGCAACGCCCTCCA
AACCTCCAGCCTCCCTGTCTCCAGCTGCCTGGGCCCCGAAGGGCTTTGGTTCTTCTCTGGGTCTGATTT
TCTACTGAACTCCACCGACCAACTGCCCTAAGCCCCCAGGGCCTCCAGGGCCAGGTTGAGACCCAAA
CCCCAAAATCCAAAATTTCTTGAAGGTTTCAAGGACCGTCCAGGGGAGATGGGGAGGAGATAGGAG
TGAGTCACCTGCTCCAGAAGATGCCAGCTTCTCTCCAGGGTGTAGTTGGCTTTGCCCCACCCCTCAC
TCCCCAGGGAGCTCCGGGGACAGCTTCTCACACCCCTGTCCACCCACACAGCTGCCCTAGCTGACCCC
GAGAAGTGCTCTTGGCTGACCCCTCTGGTGTGTGGTGAAGGGCTTTCTTCCCTTCTCTGTTTCAGACC

CCCCATTTCCCGCACATGGTGTGGGGGGCTGGGGGAGGTCCAAGCAGAGTGTATTTATTATTATCGCTTT
ATGTTTTTGGTTATTGGTTTTTTTGTATAGACCAAAGCAAAGAAAATAAAAATAACACAG

SEQ ID NO:407

5 >gi|4504811|ref|NP_002221.1| junction plakoglobin, isoform 1; gamma-catenin
[Homo sapiens]

MEVMNLMEQPIKVTEWQOTYTYDSGIHSGANTCVPSVSSKGIMEEDEACGRQYTLKKTTTYTQGVPPSQG
DLEYQMSTTARAKRVREAMCPGVSGEDSSLLLATQVEGQATNLQRLAEPSSLKSAIVHLINYQDDAELA
10 TRALPELTCLLNDEDPVVVTKAAMIVNQLSKKEASRRALMGSPQLVAAVVRTMQNTSDLDRTARCTTSILH
NLSHREGLLAIFKSGGIPALVRMLSSPVESVLFYAITTLHNLLLYQEGAKMAVRLADGLQKMPVLLNKN
NPKFLAITTDCLQLLAYGNQESKLIILANGGPQALVQIMRNYSYEKLLWTTSRVLKVLVSVCPSPNKAIVE
AGGMQALGKHLTSNSPRLVQONCLWTLRLNSDVATKQEGLESVLKILVNQLSVDDVNVLTCTATGTLNLTCT
NNSKNKTLVTQNSGVEALIHAILRAGDKDDITEPAVCALRHLSRHPEAEMAQNSVRLNYGIPAIVKLLN
QPNQWPLVKATIGLIRNLALCPANHAPLQEAAPVIRLVQLLVKAHQDAQRHVAAGTQQPYTDGVRMEIIV
15 EGCTGALHILARDPMNRMEIFRLNTIPLFVQLLYSSVENIQRVAAGVLCELAQDKEADAIDAEGASAPL
MELLHSRNEGATATYAAAVLFRISEDKNPDYRKRVSVELTNSLFKHDPAAWEAAQSMIPINEPYGDDMDAT
YRPMYSSDVPLDPLEMHMDMDGDYPIDTYSDLRPPYPTADHMLA

SEQ ID NO:408

20 >gi|12056467|ref|NM_021991.1| Homo sapiens junction plakoglobin (JUP),
transcript variant 2, mRNA

CGCCAGAGTCCGGAGCAGCCGCGCCGACCGCGCCGAGCTCAGTTCGCTGTCCGCGCCGGCTCCACCC
CGGCCCGACCCCGACCCGGCCCGGTGAGGCCCATACTCAGTAGCCACGATGGAGGTGATGAACCTGATG
GAGCAGCCTATCAAGGTGACTGAGTGGCAGCAGACATACCTACGACTCGGGTATCCACTCGGGCGCCA
25 ACACCTGCGTGCCCTCCGTGAGCAGCAAGGGCATCATGGAGGAGGATGAGGCCTGCGGGCGCCAGTACAC
GCTCAAGAAAACCACTTACACCCAGGGGGTCCCCCAGCCAAGGTGACCTGGAGTACCAGATGTCC
ACAACAGCCAGGGCCAAACGGGTGCGGGAGGCCATGTCCTGGTGTGTCAGGCGAGGACAGCTCGCTTC
TGCTGGCCACCCAGGTGGAGGGGAGGCCACCAACCTGACGCGACTGGCCGAGCCGTCCAGCTGCTCAA
GTGCGCCATTGTGCATCTCATCACTACCAGGACGATGCCGAGCTGGCCACTCGCGCCCTGCCGAGCTC
30 ACCAACTGCTCAACGACGAGGACCCGGTGGTGGTGACCAAGGCGGCCATGATTGTGAACAGCTGTGCA
AGAAGGAGGCGTCGCGGCGGGCCCTGATGGGCTCGCCCCAGCTGGTGGCCGCTGTCGTGCGTACCATGCA
GAATACCAGCGACCTGGACACAGCCCGCTGCACCACAGCATCCTGCACAACCTCTCCACCACCGGGAG
GGGCTGCTCGCCATCTTCAAGTCGGGTGGCATCCCTGCTCTGGTCCGCATGCTCAGCTCCCCCTGTGGAGT
CGGTCCTGTCTATGCCATCACCACGCTGCACAACCTGCTCCTGTACCAGGAGGGCGCCAAGATGGCCGT
35 CGCCTGGCCGACGGGTGCAAAAGATGGTGGCCCCCTGCTCAACAAGAACCAACCCCAAGTTCCTGGCCATC
ACCACCGACTGCCTGCAGCTCCTGGCCTACGGCAACAGGAGAGCAAGCTGATCATCCTGGCCAATGGTG
GGCCCCAGGCCCTCGTGACATCATGCGTAACCTACAGTTATGAAAAGCTGCTCTGGACCAACGATCGTGT
GCTCAAGGTGCTATCCGTGTGTCCAGCAATAAGCCTGCCATTGTGGAGGCTGGTGGGATGCAGGCCCTG
GGCAAGCACCTGACCAGCAACAGCCCCCGCCTGGTGCAAGTGCCTGTGGACCCTGCGCAACCTCTCAG
40 ATGTGGCCACCAAGCAGGAGGGCCTGGAGAGTGTGCTGAAGATTCTGGTGAATCAGCTGAGTGTGGATGA
CGTCAACGTCTCACCTGTGCCACGGGCACACTCTCAACCTGACATGCAACAACAGCAAGAACAAGACG
CTGGTGACACAGAACAGCGGTGTGGAGGCTCTCATCCATGCCATCCTGCGTGCTGGTGACAAAGGACGACA
TCACGGAGGCTGCCGTCTGCGCTCTGCGCCACCTCACTAGCCGCCACCCTGAGGCCGAGATGGCCCCAGAA
CTCTGTGCGTCTCAACTATGGCATCCCGCCATCGTGAAGCTGCTCAACCAGCCCAACCAAGTGGCCACTG
45 GTCAAGGCAACCATCGGCTTGATCAGGAATCTGGCCCTGTGCCCCAGCCAACCATGCCCGCTGCAGGAGG
CAGCGGTCTATCCCCCGCCTCGTCCAACCTGCTGGTGAAGGCCACCAGGATGCCAGCGCCACCTAGCTGC
AGGCACACAGCAGCCCTACACGGATGGTGTGAGGATGGAGGAGATTGTGGAGGGCTGCACCGGAGCACTG
CACATCCTCGCCCGGGACCCCATGAACCGCATGGAGATCTCCGGCTCAACACCATTCCCCTGTTTGTGTC
AGCTCCTGTACTCGTTCGGTGGAGAACATCCAGCGCGTGGCTGCCGGGGTGTGCTGTGTGAGCTGGCCAGGA
50 CAAGGAGGCGGGCCGACGCCATTGATGCAGAGGGGGCCTCGGCCCACTCATGGAGTTGCTGCACTCCCGC
AACGAGGGCACTGCCACCTACGCTGCTGCCGTCTGTTCGGCATCTCCGAGGACAAGAACCAGACTACC
GGAAGCGCGTGTCCGTGGAGCTACCAACTCCCTCTTCAAGCATGACCCGGCTGCCTGGGAGGCTGCCCA
GAGCATGATTCCCATCAATGAGCCCTATGGAGATGAGATGGACCTACCGCCCCATGTAATCCAGC
GATGTGCCCTTGACCCGCTGGAGATGCACATGGACATGGATGGAGACTACCCATCGACACCTACAGCG
55 ACGGCCTCAGGCCCCCGTACCCCACTGCAGACCACATGCTGGCCTAGGCGGCTGGCCCCAGTGGCGGTT
CTCATCTGAGAGGCTCTCCGTGCAGGCGATGGGGCAAGACAGAAAAGTGCCTGAGCTGGGGAAGCCGGGG
TGTAACCTTCTGTGTCACCTGCGCCTCCAGAGGTCTCCGTAGGGTCTTTCTTGGGATAGTGTCTGCT
CCTGCTTTTCTGTCTGCTGGGCATGGGTCCAGGGCCTGACACCCCTCCCCGCCCCCTGTGGCCCTGGCCACT
AAAGCTTCAGACTCAAGTACCCATTCTGTTTTCCCCCAGCAACGCCCTCCAAACCTCCAGCCTCCCTGT
60 CTCCAGCTGCCTGGGCCCCGGAAGGGCTTTGGTTCTTCTCTGGGTCTGATTTTCTCACTGAACCTCCACCG
ACCAACTGCCCTAAGCCCCCAGGGCCTCCAGGGCCAGGTTCCGAGACCCAAACCCCAAAATCCAAACT

TCTCTTTGAAAAGTTTCAGGGACCGTCCAGGGGAGATGGGGAGGAGATATGGAGTGAGTCACCTGCTCCAGA
AGATGCCAGCTTCTCTCTCCAGGGTGCTTAGTTGGCTTTGCCACCCCTCACTCCCCAGGGAGCTCCGGG
GACAGCTTTCCTCACACCCCTGTCCACCCACACAGCTGCCCTAGCTGACCCCGAGAAGTGCTCTTGGCTG
ACCCCTCTGGTGTGTGGTGAGGGGCTTTCTCTTCCCTTCTCTGTTTCAGACCCCCCATTTCCCGCACAT
5 GGTGTGGGGGGCTGGGGGAGGTCCAAGCAGAGTGTTTTATTATTATCGCTTTATGTTTTTGGTTATTGGT
TTTTTTGTATAGACCAAAGCAAAGAAAATAAAAATAACACAG

SEQ ID NO:409

>gi|12056468|ref|NP_068831.1| junction plakoglobin, isoform 1; gamma-
catenin [Homo sapiens]
MEVMNLMEQPIKVTEWQQTYTYDSGIHSGANTCVPSVSSKGIIMEEDEACGRQYTLKKTTTYTQGVPPSQG
DLEYQMSTTARAKRVREAMCPGVSGEDSSLLLATQVEGQATNLQRLAEPQLKSAIVHLINYQDDAELA
TRALPELTKLNDDEDPVVVTKAAMIVNQLSKKEASRRALMGSPQLVAAVVRTMQNTSDLD TARCTTSILH
NLSHHREGLLAIFKSGGIPALVRMLSSPVESVLFYAITTLHNLLLYQEGAKMAVRLADGLQKMPVLLNKN
15 NPKFLAITTDCLQLLAYGNQESKLIILANGGPQALVQIMRNYSYEKLLWTTSRVLKVL SVCPSNKP AIVE
AGGMQALGKHLTSNSPRLVQNCLWLRLNLSDVATKQEGLESVLKILVNQLSVDDVNVLT CATGTLSNLTC
NNSKNKTLVTONSGVEALIHAILRAGDKDDITEPAVCALRHLSRHPEAEMAQNSVRLNYGIPAIVKLLN
QPNQWPLVKATIGLIRNLALCPANHAPLQEA AVIPRLVQLLVKAHQDAQRHVAAGTQQPYTDGVRMEEIV
EGCTGALHILARDPMNRMEIFRLNTIPLFVQLLYSSVENIQRVAAGVLC ELAQDKEAADAIDAEGASAPL
20 MELLHSRNEGATYAAAVLFRISEDKNPDYRKRVSVELTNSL FKHPAAWEAAQSMIP INEPYGDDMDAT
YRPMYSSDVP LDPLEMHMDMDGDYPIDTYS DGLRPPYPTADHMLA

SEQ ID NO:410

Plasminogen activator inhibitor 1 (SERPINE1; PAI-1)

>GL3_6E_1_M13F Direction: N/A

AGGCANGNTACAAGTGTGATGGCAATGTGACTGGAACAGAAATAGTTTCTATCCAGGCACATCAATAAGCTCCCT
GTAATGCCCCGTAGATTCCATCCTGCAAAAGTGGCCCTCAGTTGGCGACACCCACGTGTCTGCATGATCCCCGCA
GTTGTGTGCTACGCACGTAGGACGAGCGTTGGAATGCAAGGTTGCCCANATCGGTCCCGCANGGCTTCACCGTC
TGTTTGGATGACCTATCTAAGGGGAAGTTTCGTGCTTCTAAACGATTCTTCTTCTCCAAAGGGTCTCAAGGGT
30 CGGTAGACTATGGGAGCGATTTTGACCTTAAAAGGTTTNCCTTCCCAAGNGAGAGAAGGGCCCAACGGGAGTCCT
TGGGACGAAGCGAGGGAGTCCTGTGCTTTTACTCAACATCTGTCTTAATCAATAGAGTGNCAAGATGCTCTATCG
TGCTTTCCCTTCAGAGAAGTGAGGTTCAATTGTAACAGCATAAC

SEQ ID NO:411

>GL3_37G_PCR_G3F1 Direction: N/A

TCCGAGTTTCTTCTTTCCCGATGCATCTCCAGTTTTGTCCCAGATGAAGGCGTCTTTCCCGAGGGTCAGGGGTTT
CATCACTTGCCCATGAAAAGGACTGTTCCTGTGGGGTGTGCCGGAACAACAAGAGGAAGGGTCTGTCCATGA
TGATCTCCTCGGGGGCCATGCGGGCTGAGACTATGACAGCTGTGGATGAGGAGGCCACCGTGCCACTCTCGTTCA
CCTCGATCTTCACTTTCTGCGAGCGCCTGCGCGACGTGAGAGGCTCTTGGTCTGAAAGACTCGTGAAAGTCAGCCTG
40 AAACGTGTCTGAACATGTCGGTCAATCCAGGTTCTCTAGGGGCTTCTGAGGTCGACTTCAGTCTCAGGGAGAA
CTTGGGCAGAACAGGAGGCGGGGCAGCCTGGTCATGTTGCCCTTCCAGTGGCTGATGAGCTGGGCACCTCAGAAT
GTTGGTGAGGGCAGAGAGAGGCACCTCTTTTTCATAAGGGGCAGCAATGAACATGCTGAGGGTGTCCCGTGTA
GGCAGTTCAGGATGTCGTATAATGGCCATCGGGCGTGGTGAAC

SEQ ID NO:412

>GL3_6E_1_M13R Direction: anti-sense

GACACACACAAAAGGTATGATCAGCAACTTGCTTGGGAAAGGAGCCGTGGACCAGCTGACACGGCTGGTGCTGGT
GAATGCCCTCTACTTCAACGGCCAGTGGAAAGACTCCCTTCCCGACTCCAGCACCCACCGCCGCTCTTCCACAA
ATCAGACGGCAGCACTGTCTCTGTGCCCATGATGGCTCAGACCAACAAGTTCAACTATACTGAGTTACACAGCC
50 CGATGGCCATTACTACGACATCCTGGAAC TGCCCTACCACGGGGACACCCCTCAGCATGTTTATTGCTGCCCCCTTA
TGAAAAAGAGGTGCCCTCTCTTGCCCTACCAACATTCTGAGTGCCCGAGCTCCATCAAGCCACTGGAAAGGCAAC
ATGACCAGGGGC

SEQ ID NO:413

>GL3_6E_2_M13F Direction: anti-sense

GACACACACAAAAGGTATGATCAGCAACTTGCTTGGGAAAGGAGCCGTGGACCAGCTGACACGGCTGGTGCTGGT
GAATGCCCTCTACTTCAACGGCCAGTGGAAAGACTCCCTTCCCGACTCCAGCACCCACCGCCGCTCTTCCACAA
ATCAGACGGCAGCACTGTCTCTGTGCCCATGATGGCTCAGACCAACAAGTTCAACTATACTGAGTTACACAGCC
CGATGGCCATTACTACGACATCCTGGAAC TGCCCTACCACGGGGACACCCCTCAGCATGTTTATTGCTGCCCCCTTA
60 TGAAAAAGAGGTGCCCTCTCTTGCCCTACCAACATTCTGAGTGCCCGAGCTCATCAGCCACTGGAAAGGCAACAT

GACCAGGCTGCCCCGCCTCCTGGTTCTGCCCAAGTTCTCCCTGGAGACTGAAGTCGACCTCAGGAAGCCCCCTAGA
GAACCTGGGAATGACCGACATGTTTCAGACTCGTTTCAGGCTGACTTCACGAGTCTTTTCAGACCAAGAGCCTCTCC
ACGTGCGCGCAGGCCTGAG

5 SEQ ID NO:414

>GL3_6E_3_M13F Direction: anti-sense

10 GACACACACAAAAGGTATGATCAGCAACTTGCTTGGGAAAAGGAGCCGTGGACCAGCTGACACGGCTGGTGCTGGT
GAATGCCCTCTACTTCAACGGCCAGTGAAGACTCCCTTCCCCGACTCCAGCACCCACCGCCGCTCTTCCACAA
ATCAGACGGCAGCACTGTCTCTGTGCCCATGATGGCTCAGACCAACAAGTTCAACTATACTGAGGTTACCACGCG
CCGATGGCCATTACTACGACATCCCTGGAAGTGCCTACCACGGGGACACCCTCAGCATGTTTCATTGCTGCCCCCT
T

SEQ ID NO:415

>GL3_36B_PCR_G3F1 Direction: N/A

15 AGGCAGTACAAGTGTGATGGCAATGTGACTGGAACAGAAATAGTTTCTACCAGGCACACAAAAGCTCCTGTAAGC
CCCGTAGTTCCATCCTGCAAAGGGCCTCAGTGGGAACCAAGTCTTGGCAACCCATGGGCAGAGAGACGGGTGGAAG
CAGGTGCCCCAGATGGTCCCGCAGGCGTCACCGTCTGGTTTGGAGACCTTAAGGGAGTTGTGCTTCAAACCTTCTC
TCCCAGGGGTCTCAGGTGGAGACTAGGGAGTTTGACCTAAAGGTCCCTCCAAGGAGAGGCCAAGGTCTTGGAGACA
20 GATCTGGTTTACCATCTTTTAACAAAAGGCAAATGTCTTCTCTTCCCTTCAGAAAAGAGTCATTAACACTAAAATTC
TTTCTTTCGGAGTTTCTTCTTTCCCGATGCATCTCCAGTTTGTGCCAGATGAAGGCGTCTTTCCTCCAGGGGTCA
GGGTTCCATCACTTGGCCCATGAAAAGGACTGTTCTGTGGGGTTGTGCCGACCACAAAGAGG

SEQ ID NO:416

>GL3_6E_2_M13R Direction: sense

25 AGGCAGTACAAGTGTGATGGCAATGTGACTGGAACAGAAATAGTTTCTACCAGGCACACAAAAGCTCCTGTAAGC
CCCGTAGTTCCATCCTGCAAAGGGGCTCAGTGGGAACCAAGTCTGCAGACCCGAGTGGGCAGAGAGACGGGTGG
AAGCAGGTGCCCCAGATGGTCCCGCAGGCGTCACCGTCTGGTTTGGAGACCTTAAGGGAGTTGTGCTTCAAACCTT
CTCTCCCAGGGTCTCAGGTGGAGACTAGGGAGTTTGACCTAAAGGTCCCTCCAAGGAGAGGCCAAGGTCTTGGAGA
30 CAGATCTGGTTTACCATCTTTTAACAAAAGGCAAATGTCTTCTCTTCCCTTCAGAAAGAGTCATTAACACTAAAAT
TCTTTTCTTCGGAGTTTCTTCTTTCCTCCGATGCATCTCCAGTTTGT

SEQ ID NO:417

>GL3_6E_3_M13R Direction: sense

35 CAGCTGGANCGAGCTCGGATCACTTANTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTCTCTCGGCATGGACGA
GCTGTACAAGGAGGAGGCGCCAAGGCCGTTGGCAGCGGTGGCTCCAGTGTGCTGGGTTCTTGGCGAGGCAGTACA
AGTGTGATGGCAATGTGACTGGAACAGAAATAGTTTCTACCAGGCACACAAAAGCTCCTGTAAGCCCCGTAGTTTC
CATCTGCAAAAGGGCCTCAGTGGGAACCAAGTCTGCAGACCCGAGTGGGCAGAGAGACGGGTGGAAGCAGGTGC
CCCAGATGGTCCCGCAGGCGTCACCGTCTGGTTTGGAGACCTTAAGGGAGTTGTGCTTCAAACCTTCTCTCCCAGG
40 GTCTCAGGTGGAGACTAGGGAGTTTGACCTAAAGGTCCCTCCAAGGAGAGGCCAAGGTCTTGGAGACAGATCTGGT
TTACCATCTTTTAACAAAAGGCAAATGTCTTCTCTTCCCTTCAGAAAGAGTCATTAACACTAAAATTCCTTTTCT

SEQ ID NO:418

>GL3_6E_PCR_G3F1 Direction: N/A

45 AGGCAGTACAAGTGTGATGGCAATGTGACTGGAACAGAAATAGTTTCTACCAGGCACACAAAAGCTCCTGTAAGC
CCCGTAGTTCCATCCTGCAAAGGGCCTCAGTGGGAACCAAGTCTGCAGACCCAGTGGGCAGAGAGACGGGTGGAA
GCAGGTGCCCCAGATGGTCCCGCAGGCGTCACCGTCTGGTTTGGAGACCTTAAGGGAGTTGTGCTTCAAACCTTCT
CTCCCAGGTCTCAGGTGGAGACTAGGGAGTTTGACCTAAAGGTCCCTCCAAGGAGAGGCCAAGGTCTTGGAGACA
GATCTGGTTTACCATCTTTTAACAAAAGGCAAATGTCTTCTCTTCCCTTCAGAAAGAGTCATTAACACTAAAATTC
50 TTTTCTTCGGAGTTTCTTCTTTCCTCCGATGCATCTCCCGTTTGTGCCAGATAAGGCGTCTTTCCTCCAGGGTCAGG
GGTTCCATCACTTGGCCCATGAAAAGGACTGTTCTCT

SEQ ID NO:419

>gi|10835158|ref|NM_000602.1| Homo sapiens serine (or cysteine) proteinase
inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member
55 1 (SERPINE1), mRNA

GAATTCCTGCAGCTCAGCAGCCGCCGCGCCAGAGCAGGACGAACCGCCAATCGCAAGGCACCTCTGAGA
TCAGGATGCAGATGTCTCCAGCCCTCACCTGCCTAGTCTGGGCTGGCCCTTGTCTTTGGTGAAGGGTC
TGCTGTGCACCATCCCCATCCTACGTGGCCCCACCTGGCCTCAGACTTCGGGGTGAGGGTGTTCAGCAG
60 GTGGCGCAGGCCTCCAAGGACCGCAACGTGGTTTCTCACCTATGGGGTGGCCTCGGTGTTGGCCATGC
TCCAGCTGACAACAGGAGGAGAAACCCAGCAGCAGATTCAAGCAGCTATGGGATTCAAGATTGATGACAA

GGGCATGGCCCCCGCCCTCCGGCATCTGTACAAGGAGCTCATGGGGCCATGGAACAAGGATGAGATCAGC
 ACCACAGACGCGATCTTCGTCCAGCGGGATCTGAAGCTGGTCCAGGGCTTCATGCCCCACTTCTTCAGGC
 TGTTCGGGAGCACGGTCAAGCAAGTGGACTTTTCAGAGGTGGAGAGAGCCAGATTTCATCATCAATGACTG
 5 GGTGAAGACACACACAAAAGGTATGATCAGCAACTTGCTTGGGAAAGGAGCCGTGGACCAGCTGACACGG
 CTGGTGTCTGGTGAATGCCCTCTACTTCAACGGCCAGTGGAAGACTCCCTTCCCCGACTCCAGCAGCCACC
 GCCGCTCTTCCACAAATCAGACGGCAGCACTGTCTCTGTGCCCATGATGGCTCAGACCAACAAGTTCAA
 CTATACTGAGTTACACGCCCCGATGGCCATTACTACGACATCCTGGAAC TGCCCTACCACGGGGACACC
 CTCAGCATGTTTATTGCTGCCCCCTTATGAAAAAGAGGTGCCTCTCTCTGCCCTCACCAACATTCTGAGTG
 10 CCCAGCTCATCAGCCACTGGAAGGCAACATGACCAGGCTGCCCCGCCTCCTGGTTCTGCCCAAGTTCTC
 CCTGGAGACTGAAGTCGAGCTCAGGAAGCCCCCTAGAGAACCTGGGAATGACCGACATGTTTACAGACAGTTT
 CAGGCTGACTTACAGAGTCTTTTACAGCAAGAGCCTCTCCACGTCGCGCAGGCGCTGCAGAAAGTGAAGA
 TCGAGGTGAACGAGAGTGGCAGGCTGGCCTCCTCATCCACAGCTGTCTAGTCTCAGCCCCGATGGCCCC
 CGAGGAGATCATCATGGACAGACCCCTTCTCTTTGTGGTCCGGCACAACCCACAGGAACAGTCTTTTTC
 15 ATGGGCCAAGTGATGGAACCCCTGACCCTGGGGAAAGACGCCTTCATCTGGGACAAAACCTGGAGATGCATC
 GGGAAAGAAGAACTCCGAAGAAAAGAATTTTAGTGTTAATGACTCTTTCTGAAGGAAGAGAAGACATTT
 GCCTTTTGTAAAAGATGGTAAACCAGATCTGTCTCCAAGACCTTGGCCTCTCCTTGGAGGACCTTTAGG
 TCAAACTCCCTAGTCTCCACCTGAGACCCCTGGGAGAGAAGTTTGAAGCACAACCTCCCTTAAGGTCTCCAA
 ACCAGACGGTGACGCCCTGCGGGACCATCTGGGGCACCTGCTTCCACCCGCTCTCTCTGCCCACTCGGGTCT
 20 CGAGACCTGGTTCCCACTGAGGCCCTTTCAGGATGGAAC TACGGGGCTTACAGGAGCTTTTGTGTGCCT
 GGTAGAAACTATTTCTGTTCCAGTCACATTGCCATCACTCTTGACTGCCTGCCACCGCGGAGGAGCTG
 GTGACAGGCCAAAGGCCAGTGAAGAAACACCCCTTTCATCTCAGAGTCCACTGTGGCACTGGCCACCCCT
 CCCAGTACAGGGGTGCTGCAGGTGGCAGAGTGAATGTCCCCCATCATGTGGCCCAACTCTCCTGGCCTG
 GCCATCTCCCTCCCCAGAAACAGTGTGCATGGGTTATTTTGGAGTGTAGGTGACTTGTTTACTCATTGAA
 25 GCAGATTTCTGCTTCTCTTTTATTTTATAGGAATAGAGGAAGAAATGTCAGATGCGTGCCAGCTCTTCA
 CCCCCAATCTCTTGGTGGGGAGGGGTGTACCTAAATATTTATCATATCCTTGCCCTTGAGTGCTTGTTA
 GAGAGAAAGAGAACTACTAAGGAAAATAATATTATTTAAACTCGCTCCTAGTGTTCCTTTGTGGTCTGTG
 TCACCGTATCTCAGGAAGTCCAGCCACTTGACTGGCACACACCCCTCCGGACATCCAGCGTGACGGAGCC
 CACACTGCCACCTTGTGGCCGCTGAGACCTCGCGCCCCCGCGCCCCCGCGCCCCCTCTTTTCCCTT
 30 TGATGGAATTGACCATAACAATTTCACTCCTCTCAGGGGATCAAAAGGACGGAGTGGGGGGACAGAGAC
 TCAGATGAGGACAGAGTGGTTTCCAATGTGTTCAATAGATTAGGAGCAGAAATGCAAGGGGCTGCATGA
 CCTACCAGGACAGAACTTTCCCCAATTACAGGGTGA CTACAGCCGATTGGTGACTCACTTCAATGTGT
 CATTTCCGGCTGCTGTGTGTGAGCAGTGGACACGTGAGGGGGGGGTGGGTGAGAGAGACAGGCAGCTCGG
 ATTTCAACTACCTTAGATAATATTTCTGAAAACCTACCAGCCAGAGGGTAGGGCACAAGATGGATGTAAT
 35 GCACTTTGGGAGGCCAAGGCGGGAGGATTGCTTGAGCCAGGAGTTCAAGACCAGCCTGGGCAACATAACC
 AAGACCCCGTCTCTTTAAAAATATATATATTTTAAATATACTTAAATATATATTTCTAATATCTTTAA
 TATATATATATATTTTAAAGACCAATTTATGGGAGAATTGCACACAGATGTGAAATGAATGTAATCTAAT
 AGAAGC

SEQ ID NO:420

40 >gi|10835159|ref|NP_000593.1| serine (or cysteine) proteinase inhibitor,
 clade E (nexin, plasminogen activator inhibitor type 1), member 1;
 plasminogen activator inhibitor, type I [Homo sapiens]
 MQMSPALTCLVLGLALVFGESAVHHPPSYVAHLASDFGVRVFQQVAQASKDRNVVFSPPYGVASVLAMLQ
 45 LTTGGETQQQIQAAAMGFKIDDKGMAPALRHLKELMGPWNKDEISTTDAIFVQRDLKLVLQGFMPHFRLF
 RSTVKQVDFSEVERARFIINDWVKTHTKGMISNLLGKGAVDQLTRLVLVNLALYFNGQWKTPFPDSSTHRR
 LFHKSDGSTVSVPMMAQTNKFNYTEFTTPDGHYDILELPYHGD TLSMFIAAPYEKEVPLSALTNILSAQ
 LISHWKGNMTRLRLLVLPKFSLETEVDLRKPLENLGMTDMFRQFQADFTSLSDQEPLHVAQALQKVKIE
 VNESGTVASSSTAVIVSARMAPEEIIMDRPFLFVVRHNPTGTVLFMGQVMEP

50 SEQ ID NO:421

Proteosomal subunit Y>

ACACTGCCCTCCTTCTTGAGGGTCCCAGCCTGCGATGATGATTCCCGCCATCAGGTCTTCCCGGTATCGGTAAACA
 CATCTCCTTAAAGAGACTGGCTGCTGTGTGGACCAGTGGAGGCTCATTCAAGTTCAATGCTGTGGAAACCGAGCTG
 GTAGGTGACAGCATCAGCTACTGCCTGGGTATCAGCAGCTGAGCCTGAGCGACAGCAGAAAATGCGGTGCTGAAT
 55 AGGTGTCAGCTTGTGAGTCACTCGATTGGCGATGTAGGACCCAGTGGTTGTTCTGGAGTCCGCCCCCAGAACCAC
 GCCCCCGTCAAAC TGACGGCCATGATAGTGGTCCCAGTGGAACTTCTCGGCTTTCCAGTCTGGAGTGAACGC
 CTCCGGCCCCCAAGCCGGTGTGGCCCGGCTCCCCGAGCAGCTAGTAAGGTAG

SEQ ID NO:422

Rap2B

>GL3_37F_1_M13F Direction: N/A

5 CGAAGTCTTCGATGGTTCGGGTCGTAATCCAGGATCTCCAGCACCGACGGCGACGAGTCCACCTCAATCTCCTTGCGGT
AAAAGTCTTCGATGGTTCGGGTCGTAATCCAGGATCTCCAGCACCGACGGCGACGAGTCCACCTCAATCTCCTTGCGGT
CCACGCCGCCCGAGCCCGAGCACCACTTTGTACTCTCTCATGGCTCCGTCAGCGCTCTCGCCGCGCCTGCCGC
GGCCCCGTCGGGGCTGCGCGCGGGAAAGGCTGGGCTTGCGGGCTGGACTTCTCTTCCCCCTTCTCAGCTACGCA
GGAAAAACCCAGGAGAGAGCAACCCAGAGAACGCAGGGCGGAAACACCGAAACGGAAGGCGGGCGCCAGGCC
10 GGCCCCGCGCGCCGCGGGCAAGGCGNTAATCCCTGGGGGCTAGCACGGGGTCCCGGGGGCCACGGAAGGCAAGG
TCACCGGNAGGCAGGCAAGGACANGACAGCCCCNGGNGGACGGGCAGAAAACCCGGGAANGAGGGGACAAATTGG
CCCACGGCAAGCCTTTGGCGNCACGNGAAGAAAAACACAAGGGCAACAAGGGGGNGTTAACGNCATTAAGGAA
TAACACAGCGCGGAGCCGAGAG

SEQ ID NO:423

15 >GL3_37F_4_M13F Direction: anti-sense

CGCAGCTGCGGGCATTGTCTCTCGGTTTCGCCGCCCGGGCTGCTGCTGCCGCCCGGACTGCTGCGGGGCCCGGA
CCCGCACCCAGGGATACGCTGCCGCCGTCGCCGCCCGGGCCCGGGCTCCGTTTCGGTGGTTTCCGCCCTG
CGTTCTCTGGGTTGCTCTCTCTCGGTTTTTCTGCGTAGCTGAGGAAGGGGAAGAGAAGTCCAGCCGCCAAGCC
20 CAGCCTTCCCCGGCGCAGCCCCGACGGGGCCCGGACAGGCGCGGGCGGAGAAGCGCTGACGGAGCCATGAGGA
GAGTACACAAGTGGTGGTGCTAGGGCTCGAGGCAGGAGTGGGGCAAAGTCCGCGAGACTTCAACCGTGGCAAGATA
CGGTTGGAACACGGGGCATCTTCAATCGNAGAAAAGGTACGAACCCGAACNCATCGAAAGAACTTTTTTAACCGG
CCAAGGNAGNAATGGAGAGTGGGACCTCAGGNTACGNCCCCGTCCCGGGANGCTTGGAAGAATCC

SEQ ID NO:424

25 >GL3_37F_2_M13R Direction: sense

CGAAGTCTTCGATGGTTCGGGTCGTAATCCAGGATCTCCAGCACCGACGGCGACGAGTCCACCTCAATCTCCTTGCGGT
AAAAGTCTTCGATGGTTCGGGTCGTAATCCAGGATCTCCAGCACCGACGGCGACGAGTCCACCTCAATCTCCTTGCGGT
GCCANAGACCGGACCGAGCCAGACACACCACTATATGTACTCATCTCAATGGCATCCGNCAGAGCTCATCG
30 NCCGACAGACCTGCCGNNGGACCCCGATACAGGGGGGCATTGCGAGCGCGNGGAAAAGGGCATNGGGGNCATAGG
ACAGGGACAATAGGGAACAGNTATCATATTACACACATATCGCATCAAAGCNATACGCAAAGGGAACAAAN
AACGCCCAAGAGAAGAAAGAGGCAAAACCCCCAAGGAANNGAAAACAACAGGGCCAACAAGAGGGGACCGGGGAG
GAAAAAAA

SEQ ID NO:425

35 >GL3_37F_3_M13F Direction: N/A

CGAAGTCTTCGATGGTTCGGGTCGTAATCCAGGATCTCCAGCACCGACGGCGACGAGTCCACCTCAATCTCCTTGCGGT
AAAAGTCTTCGATGGTTCGGGTCGTAATCCAGGATCTCCAGCACCGACGGCGACGAGTCCACCTCAATCTCCTTGCGGT
CCACGCCGCCCGAGCCCGAGCACCACTTTGTACTCTCTCATGGCTCCGTCAGCGCTCTCGCCGCGCCTGCCGC
40 GGCCCCGTCGGGGCTGCGCGCGGGGAAGGCTGGGCTTGCGAGGTAACCGAACGAGGCGGGCGCCGGGCCGT
GGAAAAACCCAGGAGAGAGCAACCCAGAGAACGCAGGGCGGAAACACCGAACGAGGCGGGCGCCGGGCCGT
CCGGCCGGACGANGGCAGTGCGGGGGCGAACCGAGAGGAACAAATGCCCGAGAGGCCCTTGCGGCCGAAAAA

SEQ ID NO:426

45 >GL3_37F_4_M13R Direction: sense

CGAAGTCTTCGATGGTTCGGGTCGTAATCCAGGATCTCCAGCACCGACGGCGACGAGTCCACCTCAATCTCCTTGCGGT
AAAAGTCTTCGATGGTTCGGGTCGTAATCCAGGATCTCCAGCACCGACGGCGACGAGTCCACCTCAATCTCCTTGCGGT
CCACGCCGCCCGAGCCCGAGCACCACTTTGTACTCTCTCATGGCTCCGTCAGCGCTCTCGCCGCGCCTGCCGN
GGCCCCGTCGGGGCTGCGCGCGGGGAAGGCTGGGCTTGCGGGCTGGACTTCTCTTCCCCCTTCTCAGNTACGCAG
50 GAAAAACCCAGGAGAGAGCAACCCAGGAGAACGCAGGGCGGAAAGCACGCGGAACGGAGTGCCTGGGCGCCGGGC
CGCCCCG

SEQ ID NO:427

55 >GL3_37F_PCR_G3F1 Direction: sense

CGAAGTCTTCGATGGTTCGGGTCGTAATCCAGGATCTCCAGCACCGACGGCGACGAGTCCACCTCAATCTCCTTGCGGT
AAAAGTCTTCGATGGTTCGGGTCGTAATCCAGGATCTCCAGCACCGACGGCGACGAGTCCACCTCAATCTCCTTGCGGT
CCACGCCGCCCGAGCCCGAGCACCACTTTGTACTCTCTCATGGCTCCGTCAGCGCTCTCGCCGCGCCTGCCGC
GGCCCCGTCGGGGCTGCGCGCGGGGAAGGCTGGGCTTGCGGGCTGGACTTCTCTTCCCCCTTCTCAGCTACGCAG
GAAAAACCCAGGAGAGAGCAACCCAGAGAACGCAGGGCGGAAACACCGAACGAGGCGGGCGCCGGGCCGGCC
GGCGGCGGGCGGCGAGCGTATCCCTGGGGTGCGGGTCCGGGGCCCCGAGCAGTCCGCGGGCGGCGAGCAGCAGCCGGG

CGGCGAACCGAGAGGACAATGCCCCGAGCCTGCGCCGAGAACCCAGCACAGTGGTTAGATAGATAAAGCGGCGC
TCGACTAGTCTGAGGTCTGATACTCACTGACTGAA

SEQ ID NO:428

5 >gi|16159529|ref|XM_003032.2| Homo sapiens RAP2B, member of RAS oncogene
family (RAP2B), mRNA
GCACGTTCCGCGGGGACTCATGCCACGCGCTCCCGGCCCCGACGCGCAATTAGCAGCCACCTCCGCAGCC
CGCCGCCACCGCCTCCCTGCCCTCCCGGGCTGCCGAGCTAGGAGCTCCAGCCGTCGCCTCGCGCAGGCT
10 GCGGGCATTGTCTCTCGGTTTCGCCGCCCGGGCTGCTGCTGCCGCCGCGGACTGCTGCGGGGCCCGGACC
CGCACCCCAGGGATACGCTGCCGCCGCCCGGGCCGGCCCGGCGCCCGGCTCCGTTCCGTTGGTTTCCGC
CCTGCGTTCTCTGGGTTGCTCTCTCTGGGTTTTTCTGCGTAGCTGAGGAAGGGGAAGAGAAGTCCAGC
CGCCAAGCCCAGCCTTCCCCGGCGCGAGCCCCGACGGGGCCGCGGCAGGCGCGGCAGAGCGCTGACGG
AGCCATGAGAGAGTACAAAGTGCTGGGTGCTGGGCTCGGGCGGCGTGGGCAAGTCCGCGCTCACCCTGCAG
15 TTCGTGACGGGCTCCTTTCATCGAGAAGTACGACCCGACCATCGAAGACTTTTACCGCAAGGAGATTGAGG
TGGACTCGTCGCGCTCGGTGCTGGAGATCCTGGATACGGCGGGCACCGAGCAGTTCGCGTCCATGCGGGA
CCTGTACATCAAGAACGGCCAGGGCTTCATCTGGTCTACAGCCTCGTCAACCAGCAGAGCTTCCAGGAC
ATCAAGCCCATGCGGGACCATGATCATCCGCGTGAAGCGGTACGAGCGCGTGGCCATGATCCTGGTGGGCA
ACAAGGTGGACCTGGAGGGTGAGCGCGAGGTCTCGTACGGGGAGGGCAAGGCCCTGGCTGAGGAGTGGAG
20 CTGCCCCCTTCATGGAGACGTCGGCCAAAAACAAAGCCTCGGTAGACGAGCTATTTGCCGAGATCGTGCAG
CAGATGAACTACGCGGCGCAGCCCCACGGCGATGAGGGCTGCTGCTCGGCCTGCGTGATCCTCTGAGGCG
GCCACCGCGCGCCGGCCGCGCTCTGCGCACAAAGCCAAACGCATCCGACTCTCTAAATGTGATTTATTT
CTTGCTTTGAGATTGGAGACCACTTTGCACTGGCCAGGGTGCTTGGGAGCCCGGCTGGCCTCCGCGGCC
GGCGTCCCCTGCCTCCACCTGTGCCCCGAGGGGGTGTCGGGTCTGCCCCATCCGATACTCTGGTGGAAAT
25 GTGGCTCTTTGACGATGTACGTTTCTCCCTGATTTTGGTTGATGCATATTTCCCCGTTTAAAGTAGCCGT
TAGGGCGCAGTATCGGCAGCTTGACACCCACCAAGCAAAAGTTTCAGCCTGGAAGGAGGAGGAGGAA
GGGTGGATGAAAAGGAGGGAGAGAAGGTGGAATGGTTTTTTTTTTTTTTTTCTATTTTCTTTCTTTTT
TTTTTTTTTTTTTTTTTGGTCAACAGCCGTTTTTCTAGTTCCAAGTTTAAATACATGGAGGAAGTCCGG
GAGAACCATATGAAGGAGCAGGAGGAGAGGAAGAACTTTTTTCTCTTTTCCAGGAGTAGCTGGAAG
TTAAGATCGGGTTCTTTTCTGCCAGCTTGAAGGGCAACCCCATGACTGATTGCGATTCTGAGGATGTC
30 TATGCAAAGTTGGATTCTTGTTACAGTGATCCAATCTGAAGTATTGCACATCTGAAGTGGGACTGTAA
CACTGATGCCAATACAGTGTGGGGTGCCAGAAAGTGTCTGCTGATATTTGTGG

SEQ ID NO:429

35 >gi|11433346|ref|XP_003032.1| RAP2B, member of RAS oncogene family [Homo
sapiens]
MREYKVVVLGSGGVGKSAITVQFVTGSFIEKYDPTIEDFYRKEIEVDSSPSVLEILDAGTEQFASMRDL
YIKNGQGFILVYSLVNQQSFQDIKPMRDQIIRVKRYERVPMILVGNKVDLEGEREVSYGEGKALAEWSC
PFMETSANKASVDELFAEIVRQMNYAAQPNGDEGCCSACVIL

40 SEQ ID NO:430

Semaphorin 3F

>GL3_10.20_A_M13F Direction: anti-sense

CGGCAGAAAACAGAGCCTGCCTAACAGGCCAGCCAGTTGGTGGGGCCAGGCCAGGACCACACAGTCCCCAGAC
TCAGCTGGAAGTCTACCTGCTGGACAGCCTCCGCCAAGATCTACAGGACAAAGGGAGGGAGCAAGCCCTACTCGG
45 ATGGGGCACGGACTGTCCACCTTTTCTGATGTGTGTTGTCAGCCTGTGCTGTGGCATAGACATGGATGCGAGGAC
CACTTTGGAGACTGGGGTGGCCTCAAGAGCACACAGAGAAGGGAAGAAGGGGCCATCACAGGATGCCAGCCCCTG
CCTGGGTTGGGGGCACTCAGCCACGACCAGCCCCCTTCTGGGTATTTATTCTCTATTATTGGGGATAGGAGAAG
AGGCATCCTGCCTGGGTGGGACAGGCCCTTCAGCCCCCTTCTCCNCTCCCCGCTGGCCAGGGCAGGGCCACCCCA
CTCTACCTCCTTAGCTTTCCCTGTGCCACTTTGACTCAGAGGCTGGGAGCATAGCAGAGGGGCCAGGCCAGGCA
50 GAGCTGACGGGAGGCCCCAGCTCTGAGGGGAGGGGGTC

SEQ ID NO:431

>GL3_9.67_A_M13R Direction: anti-sense

CGGCAGCAAAACAGANCTGCCTAACAGGACCCANCCAGTTGGTGGGGCCAGGCCAGGACCACACAGTCCCCAG
55 ACTCAGCTGGAAGTCTACCTGCTGGACAGCCTCCGCCAAGATCTACAGGACAAAGGGAGGGAGCAAGTCTCTC
GGATGGGGCACGGACTGTCCACCTTTTCTGATGTGTGTTGTCAGCCTGTGCTGTGGCATAGACATGGATGCGAGG
ACCACTTTGGAGACTGGGGTGGCCTCAAGAGCACACAGAGAAGGGAAGAAGGGGCCATCACAGGATGCCAGCCCC
TGCCTGGGTTGGGGGCACTCAGCCACGACCAGCCCCCTTCTGGGTATTTATTCTCTATTATTGGGGATAGGAGA
AGAGGCATCCTGCCTGGGTGGGACAGGCCCTTCAGCCCCCTTCTCCNCTCCCCGACTGGCCAGGGCAGGGCCACCC
60 CACTCTACCTCCTNANATTCCCCTGTGC

SEQ ID NO:432

>GL3_9.67_B_M13F Direction: N/A

5 GCCAGTGTATTATTAGCAAGATGGAACCCAAAGGCGGCTGTGGCCTGGGCAGCAGAAGGCCACCAGGAGCCCCAC
CCATCTACCCAAGTACCCCCCAGAGCTAATTACATCCACACCCATCCCCCTGAAGTGGTGGACATAAGGGAGCCCT
GGGGAGCCTCTACCGGCCCCAGGCCTCTACCACGGACCCCCCTCCCCCTCAGAGCTGGGGCCTCCCGTCAGCTCTGC
CTGGGCCTGGCCCCCTCTGCTATGCTCCAGCCTCTGAGTCAAAGTGGCACAGGGAAAGCTAAGGAGGTAGAGTGG
GGTGGCCCTGCCCTGGCCAGGCGGGAGGGGAGAAGGGGCTGAAGGGGCTGTCCACCCAGGCAGGATGCCTCTT
10 CTCCTATCCCCAATAAATAGAGAATAAATACCCAGGAAGGGGCTGGTCTGGCTGAGTGCCCCCAACCCAGGCAG
GGGCTGGCATCCTGTGATGGCCCCCTTCTTCCCTTCTCTGTGTGCTCTTGAGGCCACCCAGTCTCCAAAGTGGTC
CTCGCATCCATGTCTATGCCACAGCACAGGCTGACAACACACATCANAAAAGGTGGACAGTCCGTGCCCCATCCG
AGTAGGGCTTGCTCCCTCCCTTTGTCTGTAGATCTTGGCGGANGCTGTCCAGCAGGTANACTTCCAGCTGANTC
TGGGACTGTGTGGTCTTGGCCTGGCCACCAACTGGNTGGCCTGGTTAGGCAGGTCTGTTTTCTTGC

15 SEQ ID NO:433

>GL3_9.67_B_M13R Direction: anti-sense

CGGCAGAAAACAGAGCCTGCCTAACCAGGCCAGCCAGTTGGTGGGGCCAGGCCAGGACCACACAGTCCCCAGAC
TCAGCTGGAAGTCTACCTGCTGGACAGCCTCCGCCAAGATCTACAGGACAAAGGGAGGGAGCAAGCCCTACTCGG
20 ATGGGGCACGGACTGTCCACCTTTTCTGATGTGTGTTGTGAGCCTGTGCTGTGGCATAGACATGGATGCGAGGAC
CACTTTGGAGACTGGGGTGGCCTCAAGAGCACACAGAGAAGGGAAGAAGGGGCCATCACAGGATGCCAGCCCCCTG
CCTGGGTTGGGGGCACTCAGCCACGACCAGCCCCCTTCTGGGTATTTATTCTCTATTTATTGGGGATAGGAGAAG
AGGCATCCTGCTGGGTGGGACAGCCCCCTTACGCCCCCTTCTCCNCTCCCCGCTGGCCAGGGCAGGGCCACCCCA
CTCTACCTCCTTAGCTTTCCCTGTGCCACTTTGACTCAGAGGCTGGGAGCATAGCAGAGGGGGCAGGGCCAGGCA
GAGCTGACGGGAGGCCCCAGCTCTGAGGGGAGGGGCTCCGTGGTAGAGGCCTGGGGCCGGTAGAGGCTCCCCAGG
25 GCTCCCTTATGTCCACCACTTCAGGGGATGGGTGTG

SEQ ID NO:434

>GL3_9.67_C_M13F Direction: N/A

30 GCCAGTGTATTATTAGCAAGATGGAACCCAAAGGCGGCTGTGGCCTGGGCAGCAGAAGGCCACCAGGAGCCCCAC
CCATCTACCCAAGTACCCCCCAGAGCTAATTACATCCACACCCATCCCCCTGAAGTGGTGGACATAAGGGAGCCCT
GGGGAGCCTCTACCGGCCCCAGGCCTCTACCACGGACCCCCCTCCCCCTCAGAGCTGGGGCCTCCCGTCAGCTCTGC
CTGGGCCTGGCCCCCTCTGCTATGCTCCAGCCTCTGAGTCAAAGTGGCACAGGGAAAGCTAAGGAGGTAGAGTGG
GGTGGCCCTGCCCTGGCCAGGCGGGAGGGGAGAAGGGGCTGAAGGGGCTGTCCACCCAGGCAGGATGCCTCTT
CTCCTATCCCCAATAAATAGAGAATAAATACCCAGGAAGGGGCTGGTCTGGCTGAGTGCCCCCAACCCAGGCAG
35 GGGCTGGCATCCTGTGATGGCCCCCTTCTTCCCTTCTCTGTGTGCTCTTGAGGNCACCCAGTCTCCAAAGTGGTC
CTCGCATCCATGTCTATGCCACAGCACAGGCTGACAACACACATCANAAAAGGTGGACAGTCCGTGCCCCATCCG
ANTAGGGCTTGCTCCCTCCCTTTGTCTGTAAATC

SEQ ID NO:435

40 >GL3_9.67_D_M13R Direction: sense

GCCAGTGTATTATTAGCAAGATGGAACCCAAAGGCGGCTGTGGCCTGGGCAGCAGAAGGCCACCAGGAGCCCCAC
CCATCTACCCAAGTACCCCCCAGAGCTAATTACATCCACACCCATCCCCCTGAAGTGGTGGACATAAGGGAGCCCT
GGGGAGCCTCTACCGGCCCCAGGCCTCTACCACGGACCCCCCTCCCCCTCAGAGCTGGGGCCTCCCGTCAGCTCTGC
45 CTGGGCCTGGCCCCCTCTGCTATGCTCCAGCCTCTGAGTCAAAGTGGCACAGGGAAAGCTAANGAGGTAGAGTGG
GGTGGCCCTGCCCTGGCCAGGCGGGAGGGGAGAAGGGGCTGAAGGGGCTGTCCACCCAGGCAGGATGCCTCTT
CTCCTATCCCCAATAAATAGAGAATAAATACCCAGGAAGGGGCTGGTCTGGCTGAGTGCCCCCAACCCANGCAG
GGGCTGGCATCCTGTGATGGCCCCCTTCTTCCCTTCTCTGTGTGCTCTTGAGGCCACCCAGTCTCCAAAGTGGTC
CTCGCATCCATGTCTATGCCACAGCACAGGCTGACAACACACATCANAAAAGGTGGACAG

50 SEQ ID NO:436

>GL3_9.67_PCR_G3F1 Direction: N/A

GCCAGTGTATTATTAGCAAGATGGAACCCAAAGGCGGCTGTGGCCTGGGCAGCAGAAGGCCACCAGGAGCCCCAC
CCATCTACCCAAGTACCCCCCAGAGCTAATTACATCCACACCCATCCCCCTGAAGTGGTGGACATAAGGGAGCCCT
GGGGAGCCTCTACCGGCCCCAGGCCTCTACCACGGACCCCCCTCCCCCTCAGAGCTGGGGCCTCCCGTCAGCTCTGC
55 CTGGGCCTGGCCCCCTCTGCTATGCTCCAGCCTCTGAGTCAAAGTGGCACAGGGAAAGCTAAGGAGGTAGAGTGG
GGTGGCCCTGCCCTGGCCAGGCGGGAGGGGAGAAGGGGCTGAAGGGGCTGTCCACCCAGGCAGGATGCCTCTT
CTCCTATCCCCAATAAATAGAGAATAAATACCCAGGAAGGGGCTGGTCTGGCTGAGTGCCCCCAACCCAGGCAG
GGGCTGGCATCCTGTGATGGCCCCCTTCTTCCCTTCTCTGTGTGCTCTTGAGGCCACCCAGTCTCCAAAGTGGTC
CTCGCATCCATGTCTATGCCACAGCACAGGCTGACAACACACATCAGAAAAGGTGGACAGTCCGTGCCCCATCCG

AGTAGGGCTTGCTCCCTCCCTTTTGTCTGTAGATCTTGGCGGAGGCTGTCCAGCAGGTAGACTTCCAGCTGAGT
CTGGGGACTGTGTGGTCTGGCCTGGCCCCACCAACTGGCTGGGCCTGGTTAGGCAGGCTCTGTTTTCTGCCG

SEQ ID NO:437

5 >gi|1061350|gb|U38276.1|HSU38276 Human semaphorin III family homolog mRNA,
complete cds
CTGACTGGTGTCTCCCTCTCTTTTCATCTTGGGCTGTCTGCATGTGTCTCATTTCCCCACTCTCTCCTGTGC
CTCCCCTCTACTCTAATAATCAGGTCCAGGTTTCTCTGTACTGGGAGAAGACCTGTGGCTGGAGCAGGCA
10 GGGATGCACCCTATCTGTTCCCCATTCTCCAGGTGGGAGGGAGAAGGAGTAACCCACTTTATTGGCCAC
AGATGCAGGGGAGAAAGGAGAAAGCATGCTGGGAGCTGGAAAGAGCCCTAAGATCACCTGGTTTTCTAGAG
AGTGGAGCCTGCTTCTGCTAGGCCCTCCCACAATGCTTGTGCGCCGCTCTTCTCTCTGGGCTTCCCT
ACTGACTGGGGCCTGGCCATCCTTCCCTACCCAGGACCACCTCCCGGCCACGCCCCGGGTACGGCTCTCA
TTCAAAGAGCTGAAGGCCACAGGCACCGCCCACTTCTTCAACTTCTGTCTCAACACAACCGACTACCGAA
15 TCTTGCTCAAGGACGAGGACCAAGCATGCTGGGAGCTGGGAGCAAGGACTACGTGCTGTCCCTGGACCT
GCACGACATCAACCGCGAGCCCTCATTATACACTGGGCAGCCTCCCCACAGCGCATCGAGGAATGCGTG
CTCTCAGGCAAGGATGTCAACGGCGAGTGTGGGAACTTCGTGAGGCTCATCCAGCCCTGGAACCGAACAC
ACCTGTATGTGTGCGGGACAGGTGCCTACAACCCCATGTGCACCTATGTGAACCGCGGACGCCGCGCCCA
GGATTACATCTTCTACCTGGAGCCTGAGCGACTCGAGTCAGGGAAGGGCAAGTGTCCGTACGATCCCAAG
20 CTGGACACAGCATCGGCCCTCATCAATGAGGAGCTCTATGCTGGTGTGTACATCGATTTTATGGGCACTG
ATGCAGCCATCTTCCGCACACTTGGAAAGCAGACAGCCATGCGCACGGATCAGTACAACCTCCCGGTGGCT
GAACGACCCGTCGTTTCATCCATGCTGAGCTCATTCCTGACAGTGCAGGAGAATGATGATAAGCTTTACTTC
TTCTTCCGTGAGCGGTGCGGCAGAGGCGCCGAGAGCCCGCGGTGTACGCCCGCATCGGGCGCATTTGCC
TGAACGATGACGGTGGTCACTGTTGCTGGTCAACAAGTGGAGCAATTCTTGAAGGCGCGGCTCGTCTG
25 CTCTGTCCCGGCGAGGATGGCATTGAGACTCACTTTGATGAGCTCCAGGACGTGTTTGTCCAGCAGACC
CAGGACGTGAGGAACCTGTCTTTACGCTGTCTTTACCTCCTCTGGCTCCGTGTTCCGAGGCTCTGCCG
TGTGTGTCTACTCCATGGCTGATATTGCGATGGTCTTCAACGGGCCCTTTGCCACAAAGAGGGGCCCAA
CTACAGTGGATGCCCTTCTCAGGGAAGATGCCCTACCCACGGCCGGGCACGTGCCCTGGTGGAACTTC
ACGCCATCTATGAAGTCCACCAAGGATTTATCTGATGAGGTGATCAACTTCATGCGCAGCCACCCACTCA
30 TGTACCAGGCCGTGTACCTCTGCGAGCGCGGCCCTGGTAGTCCGCACAGGTGCTCCCTACCGCCTTAC
CACTATTGCCGTGGACCAGGTGGATTACGCCGACGGGCGCTATGAGGTGCTTTTCTGGGCACAGACCGC
GGGACAGTGCAGAAAGGTCAATTGTGCTGCCAAGGATGACCAAGAGATGGAGGAGCTCATGTGGAGGAGG
TGGAGTCTTTCAAGGATCAGCACCCGTCAAGACCATGACCATCTCTTCTAAGAGGCAACAACCTCTACGT
GGCGTCAGCCGTGGGTGTACACACCTGAGCCTGCACCGCTGCCAGGCGTATGGGGCTGCCTGTGCTGAC
35 TGCTGCCTTGCCCGGACCCCTTACTGTGCTGGGATGGCCAGGCCTGCTCCCGCTATACAGCATCTCCA
AGAGGCGGAGCCGCGGCAGGACGTCCGGCACGGAAACCCCATCAGGCAGTGCCGTGGGTTCAACTCCAA
TGCCAACAAGAATGCCGTGGAGTCTGTGCGATGCGGTGGCCGGCAGCGCAGCCTTCCTTGAGTGCCAG
CCCCGCTCGCCCCAAGCCACTGTTAAGTGGCTGTTCCAGCGAGATCCTGGTGACCGGCGCGGAGAGATT
GTGCAGAGGACCGCTTCCTGCGCACAGAGCAGGGCTTGTGCTCCGTGCACTGCAGCTCAGCGATCGTGG
40 CCTCTACTCCTGCACAGCCACTGAGAACAACTTTAAGCACGTCGTCACACGAGTGCAGCTGCATGTA
GGCCGGGACGCGCTCCATGCTGCCCTCTTCCACCACTGTCCATGAGCGCCCCGCCACCCCGAGCGCAG
GCCCCCAACGCTCCTTACCAGGAGTTAGCCAGCTGTGTCGCCAGCCAGAAGTGGGCCTCATCCACCA
GTACTGCCAGGGTTACTGCGCCATGTGCCCCCAGCCCCAGGGAGGCTCCAGGGGCACCCCGGTCTCCT
GAGCCCCAGGACAGAAAAAGCCCCGGAACCGCGGCACCACCTCCGGACACATGAGGCCAGCTGCCTG
45 TTCTGCCATGGGCAGCCTAGCCCTTGTCCCTTTTAATATAAAAGACATATATATATATATATATATAT
ATATAAAATATCTATATCTATACACACCCTGCCCTGCAAAGACAGTATTTATTGGTGGGTTGAATAT
AGCCTGCCTCAGTGGCAGCATCCTCCAAAACCTTAGACCCATGCTGGTTCAGAGACGGCAGAAAAACAGAGCC
TGCCCTAACAGGCCAGCCAGTTAGTGGGGCCAGGCCAGGACCACACAGTCCCCAGACTCAGCTGGAGTCT
ACCTGCTGGACAGGCCTTGCGCCAAGATCTACAGGACAAAGGGAGGGAGCAAGCCCTACTCGGATGGGGC
50 ACGGACCGTCCACCTTTCTGATGTGTGTTGTGAGCCTGTGCTGTGGCATAGACATGGATGCGAGGACCAC
TTTGGAGACTGGGGTGGCCTCAAGAGCACAGAGAGAAGGGGAAGAAGGGGCCATCACAGGATGCCACCCCT
GCCTTGGGTTGGGGGCACTCTGCCACGACCAGCCCCCTTCTGCGGTATTTATTCTCTATTTATTGGGGATA
GGAGAAGAGGCATCCTGCCTGGGTGGGACAGCCCCCTTGGGCCCCCTTCTCCCTCCCCGCTGGCCAGGGA
GGGCCACCCCACTCTACCTCCTTAGCTTTCCCTGTGCCACTTTGACTCAGAGGCTGGGAGCATAGCAGAG
GCCAGGCCAGGACAGCTGACGGGAGGCCCCAGCTCTGAGGGGAGGGGGTCCGTGGTAGAGCCTGGGGC
55 AGGTAGAGGCTCCCCAGGGCTCCCTTATGTCCACCACCTCAGGGGATGGGTGTGGATGTAATTAGCTCTG
GGGGGCACTTGGGTAGATGGGTGGGGCTCCTGGTGGCCTTCTGCTGCCAGGCCACAGCCGCTTTGGG
TTCCATCTTGCTAATAAACACTGGCTTTGGAATAAAAAAAAAAAAAAAAAA

SEQ ID NO:438

60 >gi|1061351|gb|AAB18276.1| semaphorin III family homolog
MLVAGLLWASLLTGAWPSFPTQDHLPATPRVRLSFKELKATGTAHFFNFLNTTDYRILLKDEHDHDMY

VGSKDYVLSLDLHDINREPLIIHWAASPORIEECVLSGKDVNGECGNFVRLIQPNRTHLYVCGTGAYNP
MCTYVNRGRRAQDYIFYLEPERLESGKGKCPYDPKLDTASALINEELYAGVYIDFMGTDAAI FRTLGLKQT
AMRTDQYNSRWLNDPSFIHAELIPDSAENDDKLYFFFRERSAEAPQSPAVYARIGRICLNDDGGHCCLVN
KWSTFLKARLVCSVPGEDGIEHFDLQDVFFVQQTQDVNPVIYAVFTSSGVSFRGSACVYSMADIRMV
5 FNGPFAHKEGPNYQWMPFSGKMPYPRPGTCTPGGTFTPSMKSTKDYPDEVINFMRSHPLMYQAVYPLQRRP
LVVRTGAPYRLTTIAVDQVDSADGRYEVFLGLTDRGTQVKVIVLPKDDQEMEELMLEEVEVFKDPAPVKT
MTISSKRQQLYVASAVGVTHLSLHRCQAYGAACADCCCLARDPYCAWDGQACSRYTASSKRRSRRQDVRHG
NPIRQCRGFNSNANKNAVESVQYGVAGSAAFLECPQPRSPQATVKWLFQRDPGDRRREIRAEDRFLRTEQG
10 LLLRALQLSDRGLYSCTATENNFKHVTVRQLHVLGRDAVHAALFPPLSMSAPPPGAGPPTPPYQELAQ
LLAQPEVGLIHQYCGQYWRHVPPSPREAPGAPRSPEPQDQKKPRNRRHHPDPT

SEQ ID NO:439

SPARC

>gi|4507170|ref|NM_003118.1| Homo sapiens secreted protein, acidic,
15 cysteine-rich (osteonectin) (SPARC), mRNA

CGGGAGAGCGCGCTCTGCCTGCCGCTGCCTGCCCTGAGGGTTCACAGCACCATGAGGGCCTGGA
TCTTCTTTCTCTTTGCTGCGCGGAGGGCCTTGGCAGCCCCCTCAGCAAGAAGCCCTGCTGATGAGAC
AGAGGTGGTGAAGAACTGTGGCAGAGGTGACTGAGGTATCTGTGGGAGCTAATCCTGTCCAGGTGGAA
20 GTAGGAGAATTTGATGATGGTGCAGAGGAAACCGAAGAGGAGGTGGTGGCGGAAAATCCCTGCCAGAACC
ACCACTGCAAACACGGCAAGGTGTGCGAGCTGGATGAGAACAACACCCCCATGTGCGTGTGCCAGGACCC
CACCAGCTGCCAGCCCCCATTTGGCGAGTTTGAGAAGGTGTGCGAGCAATGACAACAAGACCTTCGACTCT
TCCTGCCACTTCTTTGCCACAAAGTGACCCCTGGAGGGGCACCAAGAAGGGCCACAAGCTCCACCTGGACT
ACATCGGGCCTTGCAAATACATCCCCCTTGCTGGACTCTGAGCTGACCGAATTCCTCCCTGCGCATGCG
GGACTGGCTCAAGAAGCTCCTGGTCACCCCTGTATGAGAGGGATGAGGACAACAACCTTCTGACTGAGAAG
25 CAGAAGCTGCGGGTGAAGAAGATCCATGAGAATGAGAAGCGCCTGGAGGCAGGAGACCACCCCTGGAGC
TGCTGGCCCCGGGACTTCGAGAAGAACTATAACATGTACATCTTCCCTGTACACTGGCAGTTTCGGCCAGCT
GGACCAGCACCCCATTTGACGGGTACCTCTCCACACCGAGCTGGCTCCACTGCGTGCTCCCTCATCCCC
ATGGAGCATTGCAACACCCGCTTTTTCGAGACCTGTGACCTGGACAATGACAAGTACATCGCCCTGGATG
AGTGGGCCGGCTGCTTCGGCATCAAGCAGAAGGATATCGACAAGGATCTTGTGATCTAAATCCACTCCTT
30 CCACAGTACCGGATTCTCTCTTTAACCCTCCCCCTTCGTGTTTTCCCCCAATGTTTAAAATGTTTGGATGGT
TTGTTGTTCTGCTTGGAGACAAGGTGCTAATAGATTTAAGTGAATACATTAACGGTGTCTAAAATGAA
AATTCTAACCCAAGACATGACATTCTTAGCTGTAACCTTAATTAAGGCCTTTTCCACACGCATTAATA
GTCCCATTCTTCTTGGCATTGTAGCTTTGCCATTGCTTCTTCTTTAGTGTTCTGTTTGAACCTAATACT
35 CTGGGCTCTGCCTTAAACACACATTGCAGCTTCAACTTTTCTTCTTTAGTGTTCTGTTTGAACCTAATACT
TACCGAGTCAGACTTTGTGTTTCAATTTCAATTTAGGGTCTTGGCTGCCTGTGGGCTTCCCGAGTGGCCTG
GAGGTGGGCAAAGGGAAGTAACAGACACACGATGTTGTCAAGGATGGTTTTGGGACTAGAGGCTCAGTGG
TGGGAGAGATCCCTGCAGAATCCACCAACCAGAACGTGGTTTTGCCTGAGGCTGTAAGTGAAGAGAAAGATT
CTGGGGCTGTCTTATGAAAATATAGACATTCTCACATAAGCCAGTTTCATCACCATTCTCTCTTTACCT
TTCACTGACGTTTCTTTTACATTAGGCTGTTGGTTCAAACTTTTGGGAGCACGGACTGTCAGTTCTCTG
40 GGAAGTGCTCAGCGCATCTGCAGGGCTTCTCCTCCTCTGTCTTTTGGAGAACCAGGGCTCTTCTCAGGG
GCTCTAGGGACTGCCAGGCTGTTTCAGCCAGGAAGGCCAAAATCAAGAGTGAGATGTAGAAAGTTGTAAA
ATAGAAAAAGTGGAGTTGGTGAATCGGTTGTTCTTTCTCCTCACATTTGGATGATTGTCTAAGGTTTATGGG
CATGTTCTCTCTTTTCTTCAACCTCCCCCTTTGTTCTTCTTATTAATCAAGAGAACTTCAAAGTTAATGGG
45 ATGGTTCGGATCTCACAGGCTGAGAACTCGTTACCTCCAAGCATTTTCATGAAAAAGCTGCTTCTTATTA
TCATACAACTCTCACCATGATGTGAAGAGTTTCAAAATCTTTCAAATAAAAAGTAATGACTTAGAAA
CTGAAAAA

SEQ ID NO:440

>gi|4507171|ref|NP_003109.1| secreted protein, acidic, cysteine-rich
50 (osteonectin); Osteonectin (secreted protein, acidic, cysteine-rich) [Homo
sapiens]

MRWIFFLLCLAGRALAAPQOEALPDETEVVEETVAEVTESVSGANPVQVEVGEFDDGAEETEEVVAEN
PCQNHCKKHGKVCLENDNTPMCVCQDPTSCPAPIGEFEKVCSDNKTDFDSSCHFFATKCTLEGTKKGHK
LHLDYIGPCKYIPCLDSELTEFLMRDWLKNVLVTLYERDEDNNLLTEKQKLRVKKIHENEKRLEAGD
55 HPVELLARDFEKNYNMYIFPVHWQFGQLDQHPIDGYLSHTELAPLRAPLIPMEHCTTRFFETCDLDNDKY
IALDEWAGCFGIKQDKIDKDLVI

SEQ ID NO:441

ssDNA binding protein-1

>GL1_4_1_G3F1 Direction: sense

CTGGCAGACCCTGGCCCAATCCTAACAGTGCTAACTCAATTCCATACTCCTCCTCATCACCTGGCTACCTATGTG
GGACCCCTGGNTGGTGGCGGTCTCTCAGGAACACCCATTATGCCAGTCCCGCAGATTCAACAAATTCCAGTGA
CAACATCTACACAATGATTAATCCAGTGCCGCCTGGAGGCAGCCGGTCCAACCTCCCGATGGGTCCCGGCTCGAA
CGGTCCGATGGGCGGCATGGGTGGCATGGAGCCACACCACATGAATGGATCATTAGGGTCAGGCGACATAGACGG
NCTTCCAAAAAATTCTCCTAACACATAAGTGGCATTAGCAATCCTCCAGGCACCCCTCGAGATGACGGCGAG

SEQ ID NO:442

>GL1_4_1_G3R1 Direction: anti-sense

CTCGCCGTCATCTCGAGGGGTGCCTGGAGGATTGCTAATGCCACTTATGTTGTTAGGAGAATTTTTTGGAAGTCC
GTCTATGTCGCCTGACCCTAATGATCCATTTCATGTGGTGTGGCTCCATGCCACCCATGCCGCCCATCGGACCGTT
CGAGCCGGGACCCATCGGGAAGTTGGACCGGTGCCTCCAGGCGGCACTGGATTAATCATTTGTGTAGATGTTGTC
ACTGGAATTTGTTGAATCTGCGGGACTGGGCATAATGGGTGTTCTTGGAGGACCGCCACCACCAGGGGGTCCCAC
ATAGGTACCAGGTGATGAGGAGGAGTATGGAATTGAGTTAGCACTGTTAGGATTGGGCCAGGGTCTGCCAG

SEQ ID NO:443

>GL1_4_2_G3F1 Direction: sense

CTGGCAGACCCTGGCCCAATCCTAACAGTGCTAACTCAATTCCATACTCCTCCTCATCACCTGGTACCTATGTGG
GACCCCTGGCTGGTGGCGGTCTCTCAGGAACACCCATTATGCCAGTCCCGCAGATTCAACAAATTCCAGTGAC
AACATCTACACAATGATTAATCCAGTGCCGCCTGGAGGCAGCCGGTCCAACCTCCCGATGGGTCCCGGCTCGAAC
GGTCCGATGGGCGGCATGGGTGGCATGGAGCCACACCACATGAATGGATCATTAGGGTCAGGCGACATAGACGGA
CTTCCAAAAAATTCTCCTAACACATAAGTGGCATTAGCAATCCTCCAGGCACCCCTCGAGATGACGGCGAG

SEQ ID NO:444

>GL1_4_2_G3R1 Direction: anti-sense

CTCGNCGATCATCTCNAGGGGTGCCTGTAAGGATNGCATAAGTGCCACTTATGTTGTTNAGGAGAATTTTTTGGAA
GTCCGTCTATGTCGCCTGACCCTAATGATCCATTTCATGTGGTGTGGCTCCATGCCACCCATGCCGCCCATCGGAC
CGTTTCGAGCCGGGACCCATCGGGAAGTTGGACCGGTGCCTCCAGGCGGCACTGGATTAATCATTTGTGTAGATGT
TGTCACTGGAATTTGTTGAATCTGCGGGACTGGGCATAATGGGTGTTCTTGGAGGACCGCCACCACCAGGGGGTTC
CCACATAGGTACCAGGTGATGAGGAGGAGTATGGAATTGAGTTAGCACTGTTAGGATTGGGCCAGGGTCTGCCAG

SEQ ID NO:445

>gi|21314683|ref|NM_018070.2| Homo sapiens single stranded DNA binding protein 3 (SSBP3), mRNA

CCGCGCCCGGGGCTGACGCTTTGACAGCTGGAAAGAGCGCGGAGCCAGCGCCTGGGGGGGAGGGAGGGGA
GCGCGGCGAGGAGAGCGCCAGCGAGCGAGAGCGAGCGAGCGCGGGGAGGGGGCCGGGAGCGAGGGGC
AGCTCGGGAGAGCCGGAGCGGTAGCGGCGGCGGCGGCGGCGGCGGCGGAGGCTCGGCGCCCTCTTCCCTGC
AAACCATGTTTGGCAAAGGCAAAGGCTCGGCGGTGCCCTCGGATGGGCAGGCTCGGGAAAAGTTAGCTTT
ATACGTCTACGAATATTTACTGCACGTAGGAGCACAGAAATCTGCACAGACCTTCTTATCGGAGATTCTGA
TGGGAAAAAACATCACGTTGGGAGAACCGCCTGGGTTTTTGCACCTCGTGGTGGTGTGTATTTTGGGACC
TTTACTGTGCAGCTCCTGAAAGGAGAGACACTTGTGAACATTCAAGTGAAGCAAAAGCCTTTTCATGATTA
TAGTGCAGCAGCTGCCCCGAGCCCCGTGCTTGGCAACATTCCCCCAACGATGGGATGCCGGGAGGCCCC
ATCCCGCCAGGTTTCTTTTCAGGGTCTCCGGGGTCACAGCCCTCGCCGCACGCACAGCCTCCACCTCACA
ATCCTAGCAGCATGATGGGACCCACAGTCAGCCTCCGGGAGGAGTTCTTGGGACACAGCCATTGCTGCC
CAATTTCTATGGATCCACACGACAACAAGGCCACCCCAACATGGGAGGATCAATGCAGAGAATGAACCCCT
CCCCGAGGCATGGGGCCCATGGGTCCCGGCCACAGAATTACGGCAGCGGCATGAGACCACCACCCAACT
CCCTCGGCCCCCGCATGCCCGGGATTAACATGGGCCCCGGGAGCTGGCAGACCCTGGCCCAATCCTAACAG
TGCTAACTCAATTCCATACTCCTCCTCATCACCTGGTACCTATGTGGGACCCCTGGTGGTGGCGGTCTCT
CCAGGAACACCCATTATGCCAGTCCCGCAGATTCAACAAATTCCAGTGACAACATCTACACAATGATTA
ATCCAGTGCCGCCTGGAGGCAGCCGGTCCAACCTCCCGATGGGTCCCGGCTCGGACGGTCCGATGGGCGG
CATGGGTGGCATGGAGCCACACCACATGAATGGATCATTAGGGTCAGGCGACATAGACGGACTTCCAAAA
AATTCTCCTAACACATAAGTGGCATTAGCAATCCTCCAGGCACCCCTCGAGATGACGGCGAGCTAGGAG
GGAATCTTCTCCACTCCTTTTCAGAACGACAATTATTTCTCCAAGCATGACGATGAGTGTGTGATCCCCCT
TCTCCGAGACGCTGAGAGAGCAGGCATTGTCAGGCGGGAAGATGCCAGAAATTATGCAAGAAGTGAGGTGT
CATTATCCAGGAGCTGGTGGGGAGGGCATCTCCCTGCTCCCCCTCAACCCCTCCACCCCATCCACGCCC
CCTACCTTTCCCAATTTTAGTTTCATGCAATAAAAAGGCCAAACTTTTTATTCCATAAAACAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

SEQ ID NO:446

>gi|21314684|ref|NP_060540.2| single stranded DNA binding protein 3;
sequence-specific single-stranded-DNA-binding protein [Homo sapiens]
MFAKGKGSAPVSDGQAREKLALYVYELLHVGAKSAQTFLSEIRWEKNITLGEPPGFLHSWWCVFWDLY
CAAPERDTCESSEAKAFHDYSAAAAPSPVLGNIPPNDGMPGGPIPPGFFQGGPPGSQPSPHAQPPPHNP
SSMMGPHSQPPGGVPGTQPLLPNSMDPTRQQGHPNMGGSMQRMNPPRGMGPMGPGPQNYGSGMRPPPNLS
GPAMPGINMGPAGRPWPNPNSANSIPYSSSSPGTYVGPGGGGPPGTPIMPSPADSTNSSDNIYTMINP
VPPGGSRSNFPMPGPGSDGPMGGMGMEPHHMGSLGSGDIDGLPKNSPNNISGLSNPPGTPRDDGELGGN
FLHSFQNDNYSPTMSV

SEQ ID NO:447

SUMO-specific protease-1 (SUSP-1)

>G3_1_41_PCR_G3F1 Direction: sense

GGAGTCGTGGGCCGAGAGGAACCGGGCCCGGGAAGCGCCGTCGTCGTCGTCGCGGTCGCGTTCCCCCGGAGAGG
CCTGAGAAGCTCCGGGCCGCGGGCCTCGCTGCCCCGCCAGCCCGCGGACAGGCCCGGGCGCGCCTGGCCTGCCTTT
GTATAGCCCCGTCTGAACGTGGGAGCGCAGCCCGCTGACGGCTGAGCCCGAGGCCCGCAACCCTGCGGCGTCTA
CCCTCCTCCGGCGCGGCCCTCATCCCGCGAGCACGGCGCGGTGTGGGCCATGGATTAAGAAGGAGGCGGCGT
GGGAGGAGGAAGATGGCGGCCGCAAGAGCGCGGTAGCGCAGGGGAGATTACTTTTCTGGAAGCTTTGGCTAGA
TCAGAGTCTAAGAGAGATGGAGCCGAGAACCCAGCACAGTGGTTAGATGATAAGCGGCCGCTCGACTAGTTCT

SEQ ID NO:448

>GL3_1.41_A_M13F Direction: sense

GGGCGGTGGATGGGGAGTCGTGGGCCGAGAGGAACCGGGCCCGGGAAGCGCCGTCGTCGTCGTCGCGCGTTC
TCCCCCGGAGAGGCCTGAGAAGCTCGGGCCGCGGGCCTCGCTGCCCCGACGCCCGCGGACAGGCCCGGGCGCGCC
TGGCCTGCCTTTGTATAGGCCGTCTGAACGTGGGAGCGCAGCCCGCTGACGGCTGAGCCCGAGGCCCGCAACC
CTGCGGCGTCTACCTCCTCCGGCGCGGCCCTCATCCCGCGAGCACGGCGGCGGTGTGGGCCATGGATTAAGA
AGGAGGCGCGTGGGAGGAGGAAGATGGCGGCCGGAAGAGCGGCGGTAGCGCAGGGGAGATTACTTTTCTGGAA
GCTTTGGCTAGATCAGAGTCTAAGAGAGATGGAGCCGCAANAACCAGCACAGTGGTTAGATAGATAAAGCGGCCGC
TCGACTAGTCTGAGGTCTGATACTCACTGACTGTCTGTAAGGCGGAATTCCAGCACACTGGCGGCCGTTACTAGTG
GATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTTGTGAAATTGTTATCCGCTCAC
AATCCACACAACA

SEQ ID NO:449

>GL3_1.41_A_M13R Direction: anti-sense

ATCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTACGACAGTCAGTGAGTATCAGACCTCAGACTAGTC
GAGCGGCCGCTTTATCTATCTAACCAC'TGTGCTGGTTCTGCGGCTCCATCTCTCTTAGACTCTGATCTAGCCAAA
GCTTCCAGAAAAGTAATCTCCCTGCGCTACCGCCGCTCTTGCCGGCCGCCATCTTCTCCTCCCACGCCGCTC
CTTCTTAATCCATGGCCACACCGCCGCTGCTCGCCGGGATAGGGGGCCGCGCCGAGGAGGGTAGACGCCGC
AGGGTTGCGGGCTCGGGCTCAGCCGTCAGGCGGGCTGCGCTCCACGTTTACAGACGGGCCATACAAAGGCAGGC
CAGGCGCGCCCGGCCCTGTCCGCGGGCTGGCGGGCAGCGAGGCCCGCGGCCCGAGCTTCTCAGGCCTCTCCGGGG
GAACGCGACCGGCGACGACGACGCGGCTTCCCGGGCCCGGTTCTCTCGGCCACGACTCCCCATCCACCGC
CC

SEQ ID NO:450

>GL3_1.41_B_M13F Direction: sense

GGGCGGTGGATGGGGAGTCGTGGGCCGAGAGGAACCGGGCCCGGGAAGCGCCGTCGTCGTCGTCGCGGTCGCGT
TCCCCCGGAGAGGCCTGAGAAGCTCGGGCCGCGGGCCTCGCTGCCCCGCCAGCCCGCGGACAGGCCCGGGCGCGCC
TGGCCTGCCTTTGTATAGGCCCTCTGAACGTGGGAGCGCAGCCCGCTGACGGCTGAGCCCGAGGCCCGCAACC
CTGCGGCGTCTACCTCCTCCGGCGCGGCCCTCATCCCGCGAGCACGGCGGCGGTGTGGGCCATGGATTAAGA
AGGAGGCGGCGTGGGAGGAGGAAGATGGCGGCCGGAAGAGCGGCGGTAGCGCAGGGGAGATTACTTTTCTGGAA
GCTTTGGCTAGATCAGAGTCTAAGAGAGATGGAGCCGAGAACCCAGCACAGTGGTTAGATAGATAAAGCGGCCGC
TCGACTAGTCTGAGGTCTGATACTCACTGACTGTCTGAAGGGCGAATTCAGCACACTGGCGGCCGTTACTAGTGG
ATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTTGTGAAATTGTNATCCGCTCACAA
TTCCACACAACATACGAGCCGGAAGCAT

SEQ ID NO:451

>GL3_1.41_B_M13R Direction: anti-sense

CGGATCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTTCGACAGTCAGTGAGTATCAGACCTCAGACTAG
TCGAGCGGCCGCTTTATCTATCTAACCAC'TGTGCTGGTTCTGCGGCTCCATCTCTCTTAGACTCTGATCTAGCCA
AAGCTTCCAGAAAAGTAATCTCCCTGCGCTACCGCCGCTCTTGCCGGCCGCCATCTTCTCCTCCCACGCCGCGC

TCCTTCTTAATCCATGGCCACACCGCCGCGTGCCTCGCCGGGATGAGGGGCCGCGCCGGAGGAGGGTAGACGCC
GCAGGGTTGCGGGCTCGGGCTCAGCCGTCAGGCGGGCTGCGCTCCACGTTTACAGACGGGCTATACAAAGGCAG
GCCAGGCGCGCCCGGGCTGTCCGCGGGCTGGCGGGCAGCGAGGCCCGCGGCCGAGCTTCTCAGGCCTCTCCGG
GGGAACGCGACCGGCGACGACGACGACGCGCTTCCCGGGCCCGGTTCTCTCGGCCACGACTCCCCATCCACC
GCCC

SEQ ID NO:452

>GL3_1.41_C_M13F Direction: sense

GGGCGGTGGATGGGGAGTCTGTTGGGCCGAGAGGAACCGGGCCCGGAAGCGCCGTCGTCGTCGTCGCGGTGCGGT
TCCCCCGGAGAGGCCTGAGAAGCTCGGGCCGCGGGCTCGCTGCCCGCCAGCCCGCGGACAGGCCCGGGCGCGCC
TGGCCTGCCTTTGTATAGGCCCGTCTGAACGTGGGAGCGCAGCCCGCTGACGGCTGAGCCCGAGGCCCGCAACC
CTGCGGCGTCTACCTCTCCGCGCGGCCCTCATCCCGGCGAGCACGGCGGCGGTGTGGGCCATGGATTAAGA
AGGAGGCGGCGTGGGAGGAGGAAGATGGCGGCCGCAAGAGCGGCGGTAGCGCAGGGGAGATTACTTTTCTGGAA
GCTTTGGCTAGATCAGAGTCTAAGAGAGATGGAGCCGCAAGACGACAGTGGTTAGATAGATAAAGCGGCCGC
TCGACTAGTCTGAGGTCTGATACTCACTGACTGTCTGTAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTG
GATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCC

SEQ ID NO:453

>GL3_1.41_C_M13R Direction: anti-sense

GCTCGGATCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTACGACAGTCAGTGAGTATCAGACCTCAGA
CTAGTCGAGCGGCCGCTTTATCTATCTAACCCTGTGCTGGTTCTGCGGCTCCATCTCTCTTAGACTCTGATCTA
GCCAAAGCTTCCAGAAAAGTAATCTCCCTGCGCTACCGCCGCTCTTGCCGGCCGCCATCTTCTCTCTCCACGC
CGCTCTCTTCTAATCCATGGCCACACCGCCCGGTGCTCGCCGGGATGAGGGGCCGCGCCGGAGGAGGGTAGA
CGCCGAGGGTTGCGGGCTCGGGCTCAGCCGTCAGGCGGGCTGCGCTCCACGTTTACAGACGGCCCTATACAAAG
GCAGGCCAGGCGCGCCCGGGCTGTCCGCGGGCTGGCGGGCAGCGAGGCCCGCGGCCGAGCTTCTCAGGCCTCT
CCGGGGGAACGCGACCGGCGACGACGACGACGCGCTTCCCGGGCCCGGTTCTCTCGGCCACGACTCCCCATC
CACCGCCC

SEQ ID NO:454

>GL3_1.41_D_M13F Direction: sense

GGGCGGTGGATGGGGAGTCTGTTGGGCCGAGAGGAACCGGGCCCGGAAGCGCCGTCGTCGTCGTCGCGGTGCGGT
TCCCCCGGAGAGGCCTGAGAAGCTCGGGCCGCGGGCTCGCTGCCCGCCAGCCCGCGGACAGGCCCGGGCGCGCC
TGGCCTGCCTTTGTATAGGCCCGTCTGAACGTGGGAGCGCAGCCCGCTGACGGCTGAGCCCGAGGCCCGCAACC
CTGCGGCGTCTACCTCTCTCCGCGCGGCCCTCATCCCGGCGAGCACGGCGGCGGTGTGGGCCATGGATTAAGA
AGGAGGCGGCGTGGGAGGAGGAAGATGGCGGCCGCGCCGCGGTAGCGCAGGGGAGATTACTTTTCTGGAA
GCTTTGGCTAGATCAGAGTCTAAGAGAGATGGAGCCGCAAGACGACAGTGGTTAGATAGATAAAGCGGCCGC
TCGACTAGTCTGAGGTCTGATACTCACTGACTGTCTGTAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTG
GATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCG

SEQ ID NO:455

>GL3_1.41_D_M13R Direction: anti-sense

GATCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTACGACAGTCAGTGAGTATCAGACCTCAGACTAGT
CGAGCGGCCGCTTTATCTATCTAACCCTGTGCTGGTTCTGCGGCTCCATCTCTCTTAGACTCTGATCTAGCCAA
AGCTTCCAGAAAAGTAATCTCCCTGCGCTACCGCCGCTCTTGCCGGCCGCCATCTTCTCTCTCCACGCCGCT
CCTTCTTAATCCATGGGCCACACCGCCCGGTGCTCGCCGGATGAGGGGCCGCGCCGGAGGAGGTAGACGCCG
CAGGGTTGCGGGCTCGGGCTCAGCCGTCAGGCGGGCTGCGCTCCACGTTTACAGCGGGCTATACAAAGGCAGG
CCAGGCGCGCCCGGGCTGTCCGCGGGCTGGCGGGCAGCGAGGCCCGCGGCCGAGCTTCTCAGGCCTCTCCGGG
GGAACGCGACCGGCGACGACGACGACGCGCTTCCCGGGCCCGGTTCTCTCGGCCACGACTCCCCATCCACCG
CCC

SEQ ID NO:456

>gi|7662311|ref|NM_015571.1| Homo sapiens SUMO-1-specific protease (SUSP1), mRNA

ATGGCGGCCCGCAAGAGCGGCGGTAGCGCAGGGGAGATTACTTTTCTGGAAGCTTTGGCTAGATCAGAGT
CTAAGAGAGATGGAGGTTTTTAAATAATTGGAGCTTTGATCATGAAGAAGAAAGTGAAGGAGATACAGA
TAAAGATGGGACAAATCTGCTCAGTGTGGATGAAGATGAGGATTCTGAAACCTCAAAAGGAAAAAGTTA
AATCGTCGATCTGAAATTGTTGCTAATAGCTCTGGTGAATTCATCTTGAAGACATATGTAAGACGAAACA
AGTCTGAAAGTTTTTAAACTTTGAAAGGCAACCAATTGGACTTAACATGTTGAGCAACAATAAGAAATT
GAGTGAATAATGCAAAATACGTCATTATGTTCTGGAAGTGTAGTTCATGGTAGACGTTTTTCATCATGCT
CATGCACAGATACCAGTAGTAAAAACAGCAGCCCAAGCAGTCTGGACCGAAAAGAAAGGAAAGAAATACC

5 CACCTCATGTCCAAAAAGTTGAAATTAATCCTGTAAGGTTAAGTCGGCTCCAAGGTGTTGAACGTATAAT
GAAGAAAACAGAAGAGTCCGAATCACAAGTGGAGCCTGAAATTAAGAGGAAAGTACAACAGAAACGGCAC
TGTAGTACCTATCAGCCTACTCCTCTATCTCCTGCTTCAAAAAATGTTTAACCCATTTAGAGGATT
TGCAAAGAAATTGCAGACAAGCTATTACTTTGAAATGAGTCTACTGGACCATTATTAAGAACGTCAATTCA
10 TCAGAATTCTGGAGGACAGAAGTCAAAAAACAGGATTAACAACCAAGAAGTTTTATGGCAACAATGGTG
GAAAAGGTTCCAATTGATATTATTGTGAATTGTGATGACAGTAAACACACTTATTTACAGACTAATGGAA
AAGTCATTTTACCTGGGGCAAAAAATACCCAAAAATCACAACCTTGAAAGAAAGGAAAACAAGTTTGTGAGA
CCTAAATGATCCAATCATTGTTGTCCAGTGATGATGATGATGACAACGACAGAACTAACAGAAGAGAAAGC
15 ATATCTCCTCAGCCTGCTGATTGAGCATGTTCTTCCCTGCACCATCCACTGGAAAAGTAGAAGCAGCAC
TAAATGAAAATACTTGAGAGCAGAGCGTGAACCTACGAAGCATTCCAGAAGACTCAGAGTTAAATACAGT
TACATTGCCAAGAAAAGCAAGAATGAAAGACCAGTTTGGCAATTCTATTATCAACACACCTCTGAAACGT
CGTAAAGTGTCTTCTCAAGAACCTCCAGATGCTTTAGCTTTAAGCTGCCAAAGTTCCTTTGACAGTGTCA
TTTTAACTGTGCAAGTATACGAGTAGGAACACTTCCGGCTGTTAATAGAGCCTGTAATTTTTTGTGTT
20 AGATTTTATCAAGATACAGCTAGACGAACACAGACCATTCTCTGAGAGATTATATTAATACCTCTGAT
CTAACTAAATGTGAATGGTGTAAATGTCCGAAAATTACCTGTAGTGTCTTCAAGCAATTCAGCAGTTT
ATCAAAAGCTGAGCATCCAATGCAATGAATAAGGAGGATAAAGTTTGAATGATTGTAAAGGAGTAAA
TAAATTAACAAATTTAGAAGAACAATATATAATTTTAAATTTTCAAAATGGCCTTGATCCTCCGGCAAT
ATGGTATTTGAAAGTATCATTAAATGAAATTGGTATAAAGAATAACATCTCCAATTTTTTTGCGAAAATTC
25 CCTTTGAAGAAGCTAATGGCAGACTTGTTGCCTGTACAAGAACCCTATGAAGAGAGCATCAAAGGAAGTTG
TGGGCAAAAGGAAAACAAAATTAAGAACTGTATCATTGAAATCTAAAATACAACTTAGAAGCAAAACAAGAA
TTTCAGTTTTTTTGTATGAAGAAGAAGAACTGGAGAAAACACACCATCTTCATTGGCCCAGTAGAAAAGT
TGATAGTATATCCACCACCTCCAGCTAAGGAGGCATCTCTGTTACCAATGAGGACCTGCACGTGTCTAAA
TGAAGGAGAATTTTTAAATGATGTTATTATAGACTTTTATTGAAATACTTGGTGCTTGAAAACTGAAG
30 AAGGAAGACGCTGACCGAATTCATATATTCAGTTCTTTTTTCTATAAACGCCCTTAATCAGAGAGAGAGGA
GAAATCATGAAACAACTAATCTGTCAATACAGCAAAACGGCATGGGAGAGTAAAAACATGGACCCGGCA
CGTAGATATTTTTGAGAAGGATTTTATTTTTGTACCCCTTAATGAAGCTGCACACTGGTTTTTGGCTGTT
GTTTGTCTCCCGGTTTGGAAAAACCAAGTATGAACCTAATCCTCATTACCATGAAAATGCTGTCATAC
AGAAATGTTCAACTGTAGAGGACAGTTGTATTTCTTCTCAGCCAGTGAAATGGAGAGTTGTTCCAAAA
CTCTTCTGCGCAAGCCTGTAATTAAGAAGATGCTAAACAAAAAACATTGCATAGCTGTAATTGATTCCAAT
35 CCTGGGCAGGAAGAAAGTGACCTCGTTATAAGAGAAACATATGCAGTGTAATAACAGTGTGAAAAAAA
TAAATCATACTGCGAGTGAAAAATGAAGAATTAAGAGAGAATCTACATCCCAGAAAGTTGCTGATAG
GACTAAAAGTGAGAATGGCCTACAGAATGAAAGTTTGAAGTTCCACACATCATAAGATGGCTTAAGCAAA
ATCAGACTAACTATAGCGATGAATCACCTGAAGCTGGTAAAATGCTTGAAGATGAACCTCGTCAGCTTCT
CAGAAGATCAGGATAACCAGGATGATAGCAGTGACGATGGATTCTCGCTGATGACAACCTGCAGTTCAGA
40 AATAGGACAGTGGCATTAAAGCCTACTATCTGTAAACAACCTTGATCTTACTTATGGACTCACTCCGA
GGCCCTTCTCGGTCAAATGTTGTCAAAATTTTAAAGAGAGTATTTAGAAGTGGAATGGGAAGTTAAAAAAG
GAAGCAAAAGAAGTTTTTCCAAAGATGTTATGAAGGGCTCTAATCCAAAAGTACCACAGCAAAACAACCTT
CAGTGACTGTGGTGATATGATTTGAGTATGTAGAGAGCTTTTTTGAAGATCCAATTCTCAGTTTTGAA
CTACCTATGAATTTGGCAAACTGGTTTCTCCACCAAGAATGAGAACAAAAAGAGAAGAAATCCGAAACA
45 TAATTTCTGAAGCTACAGGAAGATCAGAGCAAAAGAGAAAAGAAAGCATAAGGACACTTACTCAACAGAAGC
ACCTTTAGGCGAAGGAACAGAAACATGTGTCAATAGTATCTCAGATTGACCATTCTGTTTGTCTCAT
TCTACTTTTCAAGAACTAAATGACTTTCAAATTTGGGTATAGACAATAAAGAACTGAAGTGCTCACTACTC
AGTGATTTGGAATTTTGTATGCTGTATAAATGTGAGATAATTAATTTCCAAAGGCGTATGTATTAAGTA
AAAGTCTGTAAATATGTTAATGAGGCCAATTTTCCAGCATTTATAATTATTTTTTTCACTTGTAGGAA
50 GCTTTTGTATGATTTTTCTGTAAATAGTACCTAAAATTGCAACTTCTAAACCCAAATAAAAAAGAAAATA
TTTATAGGAGGAAATGATTAATTTGATATTCTTTAGTGAACCTGTTTAAATTCCTCAGTGGGTGTGACATA
TTTCATGGGAATATTTCAAATATCTATGGTAATATTTTGACCCTTTATATTTGTTCTAAAATAAGTCAAAA
TGTGAAAAATAATATTAATCTAAGATATTTTGAACCTAAGCATCTTTATATGCTTGTGTAACAGGAACAAA
55 GTAACAGCCTTTCAATTCATATACTGCCTTGTGTTTCAAGTGAACCCAAAGAAATGTAATAAATATTTGTAAT
TTTACACAAATATTTAAGAGGAAAGAGTATTAAGAGCAATTCAAAAAAGTAACCTTATACTACTAAAAA
AAAAATTTCTGCATATATTATCATCAAATGCATTTTGAAGACATCAAAGACTCAGGTTAAACTATTTTT
GGTAAGTGCAGCTTGAATTTCAAATATCCCGTGTACCTTTCTCTATTACAGCTTAAAGTATGCTACAAT
CTGTGTCATATAGTTAATTGATAAGCATTTTTAATCTGTGTAAACACAGGAATTTAAATAGGAATTTACT
ATTTTTTTAT

SEQ ID NO:457

>gi|7662312|ref|NP_056386.1| SUMO-1-specific protease [Homo sapiens]
MAAGKSGGSAGEITFLEALARSESKRDGGFKNNWSFDHEEESGDDTKDGTNLLSVDEDEDETSKGGKL
NRRSEIVANSSGEFILKTYVRRNKSESFKTLKGNPIGLNMLSNKKLSENMQNTSLCSGTVVHGRFRFHA
60 HAQIPVVKTAQSSSLDRKERKEYPPHVQKVEINPVRLSRLLQGVVERIMKKTEESQVEPEIKRKVQQRH
CSTYQPTPPLSPASKKCLTHLEDLQRNCRQAITLNESTGPLLRTSIHQNSGGQKSQNTGLTTKKFYGNV

EKVPIDIIIVNCDDSKHTYLQTNQKVLPGAKIPKITNLKERKTSLSLDNDPIILSSDDDDNDRTNRRES
 ISPPQADSACSSPAPSTGKVEAALNENTCRAERELRSIPEDSELNTVTLPRKARMKDQFGNSIINTPLKR
 RKVFSQEPDALALSCQSSFDVILNCRSIRVGTFLRLLIEPVIFCLDFIKIQLDEPDHDPVEIILNTSD
 LTKCEWCNVRKLPVVFLQAIIPAVYQKLSIQLOMNKEDKVWNDCKGVNKLTNLEEYIILIFQNGLDPPAN
 5 MVFESIINEIGIKNNISNFFAKIPFEEANGRLVACTRYEESIKGSCGQKENKIKTVSFESKIQLRSKQE
 FQFFDEEEETGENHTIFIGPVEKLIVYPPPPAKGGISVTNEDLHCLNEGEFLNDVIDFYLYLVLEKLEK
 KEDADRIHIFSSFFYKRLNQRRERNHETTNLSIQQKRHGRVKTWTRHVDIFEKDFIFVPLNEAAHWFLAV
 VCFPGLEKPKYEPNPHYHENAVIQKCSTVEDSCISSASEMESCSQNSSAKPVIKMLNKKHCIAVIDSN
 PGQESDPRYKRNICSVKYSVKKINHASENEEFNKGESTSQKVADRTKSENGLQNESLSSTHTDGLSK
 10 IRLNYSDESPEAGKMLEDELVDFSEDQDNQDDSSDDGFLADDNCSSSEIGQWHLKPTICKQPCILLMDSL
 GPSRSNVVKILREYLEVEWEVKKGSKRSFSKDVMMKGSNPKVPQNNFSDCGVYVLQYVESFFENPILSFE
 LPMNLANWFPFPRMRTKREEIRNIIILKLQEDQSKEKRKHKDTYSTEAPLGEGTEQCVNSISD

SEQ ID NO:458

15 Synaptophysin homologue (pantophysin)

>GL2_71_G03_G3F1 Direction: sense

ACAGGAACACAAAGACTGCAAAGGTAACATAGAATTGTGCAGAAGAAGAGTAATCGCCTATGAGGACGTAATCTT
 TCCAATTTACATCACATATGTTTACACCTGGAGGTGGCTGAAATGATGCCTCATTCACCTGAATGGATAACCAA
 AAGTAGCTGTAACAGTTTTATTCTCAGTAAGTGCAGGAGGACAATTCAGTTGAATTTCTGTTTGGCCCTTAAAC
 20 CTCCACAGGTGGCAAAGCAAAGATAGAAGCAATCCACTCGAGGACCTTGATGAAGCCGAGTGGCTCCTTGAGCG
 GGTGAGGTTGATCTGGAAGCCGCCGAGGACCCAGCGCAGTGGTTAGATAGATAAAGCGACCGCTCGACTAGTC
 TGAGGTCTGATACTCACTGACTGTCTGTAAGGGCGAATTCGTTTAAACCTGCAGGACTAGTCCCTTTAGTGAGGGT
 TAATCTGAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCACACA
 ACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTTGCGTTG
 25 CGCTCACTGGCCGCTTTCCAGTCGGGAAAACCTTGTCTGTCAGTGCATTAATGAATCGGCCAACGCCGCGGGG
 NAGAGGNCGGGTAGCGTNAATAGGGGGCGCCCTTACCCGGTTTCTCCCGCTCANTTGAATCCGCTTGCCCG
 CGGGCG

SEQ ID NO:459

30 >GL2_71_G04_G3F1 Direction: sense

ACAGGAACACAAAGACTGCAAAGGTAACATAGAATTGTGCAGAAGAAGAGTAATCGCCTATGAGGACGTAATCTT
 TCCAATTTACATCACATATGTTTACACCTGGAGGTGGCTGAAATGATGCCTCATTCACCTGAATGGATAACCAA
 AAGTAGCTGTAACAGTTTTATTCTCAGTAAGTGCAGGAGGACAATTCAGTTGAATTTCTGTTTGGCCCTTAAAC
 CTCCACAGGTGGCAAAGCAAAGATAGAAGCAATCCACTCGAGGACCTTGATGAAGCCGAGTGGCTCCTTGAGCG
 35 GGTGAGGTTGATCTGGAAGCCGCCGAGGACCCAGCGCAGTGGTTAGATAGATAAAGCGACCGCTCGACTAGTC
 TGAGGTCTGATACTCACTGACTGTCTGTAAGGGCGAATTCGTTTAAACCTGCAGGACTAGTCCCTTTANTGAGGGT
 TAATCTGAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCACACA
 ACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTTAAATGAGTTGAGCTAACTCACATTAATTTGCG
 TTTGCGCTCACTTGGGCCGCTTTCCAGGTTCCGGGGAAAACTGTCTGTCAGTTGCTTAATGAATCGGGCAACG
 40 CGCGGGGAGAGGCGGTGGGTATNG

SEQ ID NO:460

>GL2_71_G05_G3F1 Direction: sense

ACAGGAACACAAAGACTGCAAAGGTAACATAGAATTGTGCAGAAGAAGAGTAATCGCCTATGAGGACGTAATCTT
 45 TCCAATTTACATCACATATGTTTACACCTGGAGGTGGCTGAAATGATGCCTCATTCACCTGAATGGATAACCAA
 AAGTAGCTGTAACAGTTTTATTCTCAGTAAGTGCAGGAGGACAATTCAGTTGAATTTCTGTTTGGCCCTTAAAC
 CTCCACAGGTGGCAAAGCAAAGATAGAAGCAATCCACTCGAGGACCTTGATGAAGCCGAGTGGCTCCTTGAGCG
 GGTGAGGTTGATCTGGAAGCCGCCGAGGACCCAGCACAGTGGTTAGATAGATAAAGCGGCCGCTCGACTAGTC
 TGAGGTCTGATACTCACTGACTGTCTGTAAGGGCGAATTCGCGGCCGCTAAATTCAGCTTATAGTG

SEQ ID NO:461

>GL2_71_G06_G3F1 Direction: sense

ACAGGAACACAAAGACTGCAAAGGTAACATAGAATTGTGCAGAAGAAGAGTAATCGCCTATGAGGACGTAATCTT
 TCCAATTTACATCACATATGTTTACACCTGGAGGTGGCTGAAATGATGCCTCATTCACCTGAATGGATAACCAA
 55 AAGTAGCTGTAACAGTTTTATTCTCAGTAAGTGCAGGAGGACAATTCAGTTGAATTTCTGTTTGGCCCTTAAAC
 CTCCACAGGTGGCAAAGCAAAGATAGAAGCAATCCACTCGAGGACCTTGATGAAGCCGAGTGGCTCCTTGAGCG
 GGTGAGGTTGATCTGGAAGCCGCCGAGGACCCAGCACAGTGGTTAGATAGATAAAGCGGCCGCTCGACTAGTC
 TGAGGTCTGATACTCACTGACTGTCTGTAAGGGCGAATTCGCGGCCGCTAAATTCAGCTTATAGTGAGTGC
 TATTACAATTCAGTGGCCGTCGTTTACAACGTCGTGACTGGGAAAACCTTGGCGTTACCCAACCTTAATCGCCTT
 60 GCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGACCGATCGCCCTTCCCAACAGTTGCGCA

GCCTATACGTACGGCAGTTTTAAGGTTTTACNACCTATTAAAAGAGAGAGCCCGGTATCGTCTGTTGTGGATGTNC
AGAGTGATATTTATTTGGAACCCCGGGGAAAGAAT

SEQ ID NO:462

5 >gi|5803184|ref|NM_006754.1| Homo sapiens synaptophysin-like protein
(SYPL), mRNA
TGAACCGAGGCAAGGGGGCGCGGCGCAGCAGTATGGCGCCCAACATCTACTTGGTTCGCCAGCGGATCA
GTCCGACTCGGCCAGAGGATGTCCGGCTTCCAGATCAACCTCAACCCGCTCAAGGAGCCACTCGGCTTCAT
10 CAAGGTCCTCGAGTGGATTGCTTCTATCTTTGCTTTTGCCACCTGTGGAGGTTTAAAGGGCCAAACAGAA
ATTCAAGTGAATTGTCTCCTGCAGTTACTGAGAATAAACTGTTACAGCTACTTTTGGTTATCCATTCA
GGTTGAATGAGGCATCATTTTCAGCCACCTCCAGGTGTAAACATATGTGATGTAAATTGGAAAGATTACGT
CCTCATAGGCGATTACTCTTCTCTGCACAATTCTATGTTACCTTTGCAGTCTTTGTGTTCCCTGTACTGC
ATTGCTGCCCTTCTGCTTTATGTTGGCTACACGAGTCTGTATCTGGATAGTCGTAAACTTCCTATGATAG
15 ACTTTGTTGTTACACTTGTGGCACTTTTTTGTGGTTGGTGAGCACTTCAGCCTGGGCTAAAGCTCTGAC
AGATATTAATAAGCTACTGGTCACAATATTATTGATGAACCTCCGCTTGTAGAAGAAAGCAGTACTG
TGTTACTTTGGCTCTGTGACCAGTATGGGATCCCTAAATGTATCTGTGATATTTGGCTTTCTAAATATGA
TACTCTGGGGAGGAAATGCTTGGTTTGTGTACAAGGAGACCAGCCTACACAGTCCATCAAATACATCTGC
CCCTCATAGCCAAGGAGGTATTCCACCTCCTACCGGAATATAATTAAAGGGAGAAATACACTGTATGAAG
TATATGTTGATACTATGACATGTTGCCAACACCTTGAGAAGCATTATTTGTTTCTAATAAAAGTAATGGC
20 TTTGTCAATATATTGGTGGGTTTAAACTTTGCTGCTTTTTTACATAAAGCCTGTGCCTTTCTTAGAAAG
TTAAGATGTAAATGTATTCTCACATGTAAATTTGAAAGTTCAGGGGTCTATTATGAAATGGATTACACAT
TTTAAATGAACCCATAATTTTTTTCTACTAAAGCTGTTTGCCCTCCAAAGTGTTTACACCTAAGCCTAACA
TGTATCGCTCATTCAGAAAACCTGTTATATTGTATACCATAGTAGGAAGAAAAACCTTTATTTGGAATAT
ACACTACTGTAAGTTTGTACAGATCATATACCTACCACCTGTCTTTGCTTAAAGAGCCTTGATTACATAA
25 ATATGTAGGAAAAACATATTGAGTTCAAAATTTATATCTAACATTGTTTATGTTATGATTTTTTTTAAAT
TGCAAGACTAGGTGTATATTTTTTTCTGTTTTTCTAAATGACCCGTGGTACTTAATAGGTGTACTAAAA
TTGTGTTGGGAGCAGGGATTGGAATTTCTGAGAGATGTGTAGTTAATTAGTAATCTGTTTCATGAGA
TATGATCTGTTATGCTAGTGGTTTAAATAGGCTTGCTATGTAAGTAGAACGTGGCTCAACTAGATATCTTA
TATGATGGGCATTACCTCTTAGTGATATTTGTTTCTGCTCTTTGTTGCTCATGCTGTTTAAAGTGCAGG
30 CTGAGACCCAGCCTCTTTGTAAGTACAGTAAATAATCCACCGTTTTTTACAGACCCTAGTCAAAGGGTT
AAAAAAATTAAGATTGCTTTCCATGTTTGAAATTTACCATTGAGAGTCAATGAAGTTGCTATTTTGAGTT
TAGCATTGATATTGTGAAAATAAGTGCAATTTGGATTTTCATGTTTCTTAATATTCACTCTGTTTCACAA
ATGAATGATTAAGGAATTATGCATCATAAAGGAACCTAAGTGAGGTATATGATGAGTGTATTGCTTTTGC
ACACACATATAGGTATATTCTGAATACAAGCTTATTCATTTTGTCTTCTTAATCTTTTTGTTGTACAGGGA
35 TTCAGGTTTCTTATTCTTACAACATGATTGTTTATATGTGAAGCACATCTTGCTGTTGCCTTATTTTTGA
TGCTTTTATTTCATGACAAGAATTGTCAATATAAGAATGTATATCTTGCCGCAACCAATTTAATAAAGGA
GTTGAAAGAAAAA

SEQ ID NO:463

40 >gi|5803185|ref|NP_006745.1| synaptophysin-like protein; pantophysin [Homo
sapiens]
MAPNIYLVRQRISRLGQRMSSGFQINLNPLKEPLGFIKVLEWIASIFAFATCGGFKGQTEIQVNCPPAVTE
NKTVTATFGYPPRLNEASFQPPPGVNICDVNWKDYVLIGDYSSSAQFYVTFVAVFVFLYCIAALLLVGYT
45 SLYLDSRKLPMIDFVTLVATFLWLVTSAWAKALTDIKIATGHNI IDELPCKKKAVLCYFGSVTSMGS
LNVSVIFGFLNMILWGGNAWFVYKETSLSHPSNTSAPHSQGGIPPTGI

SEQ ID NO:464

Thrombospondin-1

50 >GL2_199_2_M13R Direction: anti-sense
CCAATGTAGTTAGTGCAGGATGGCAGGGCTGGAACCATTCACCACGTTGTTGTCAAGGGTGAGGAGGACACTGGTA
GAGCTGGAGCAGCCTTTGTTCTGAGGATGTCTTCTGGTGTGGTTCCAAAGACAAACCTCACATTCTGCAGCACC
CCCTGGAAAATGTCA

SEQ ID NO:465

55 >GL2_199_1_M13R Direction: sense
TGACATTTTCCAGGGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACCACACCAGAAGACATCCTCAGGAACA
AAGGCTGCTCCAGCTCTACCAGTGTNCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGGCATCC
GNACTAACTACATTGG

SEQ ID NO:466

>GL2_199_3_M13F Direction: sense

TGACATTTTCCAGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACCACACCAGAAGACATCCTCAGGAACAA
AGGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCG
CACTAACTACATTGG

SEQ ID NO:467

>GL2_66_F01_G3F1 Direction: sense

TGACATTTTCCAGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACCACACCAGAAGACATCCTCAGGAACAAA
GGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCTGGCATCCGACTA
ACTACATTGG

SEQ ID NO:468

>GL2_66_F02_G3F1 Direction: sense

TGACATTTTCCAGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACCACACCAGAAGACATCCTCAGGAACAAA
GGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCG
ACTAACTACATTGG

SEQ ID NO:469

>GL2_66_F03_G3F1 Direction: sense

TGACATTTTCCAGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACCACACCAGAAGACATCCTCAGGAACAAA
GGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCG
CACTAACTACATTGG

SEQ ID NO:470

>GL2_66_F04_G3F1 Direction: sense

TGACATTTTCCAGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACCACACCAGAAGACATCCTCAGGAACAA
AGGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCG
CACTAACTACATTGG

SEQ ID NO:471

>GL2_86_2A01_G3F1 Direction: sense

TGACATTTTCCAGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACCACACCAGAAGACATCCTCAGGAACAAA
GGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCG
CACTAACTACATTGG

SEQ ID NO:472

>GL2_86_2A02_G3F1 Direction: sense

TGACATTTTCCAGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACCACACCAGAAGACATCCTCAGGAACAAA
GGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCG
CACTAACTACATTGG

SEQ ID NO:473

>GL2_86_2A03_G3F1 Direction: sense

TGACATTTTCCAGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACCACACCAGAAGACATCCTCAGGAACAAA
GGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCG
CACTAACTACATTGG

SEQ ID NO:474

>GL2_86_2A04_G3F1 Direction: N/A

TGACATTTTCCAGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACCACACCAGAAGACATCCTCAGGAACAAA
GGCTGCTCCAGCTCTACAGCGTCTCCTCA

SEQ ID NO:475

>GL2_86_2A06_G3F1 Direction: sense

TGACATTTTCCAGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAATCACACCAGAAGACATCCTCAGGAACAAAGG
CTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCGCAC
TAACTACATTGG

SEQ ID NO:476

>GL2_86_2A07_G3F1 Direction: sense

CGACATTTTCCAGGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACACACCAGAAGACATCCTCAGGAACAA
AGGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCG
CACTAACTACATTGG

SEQ ID NO:477

>GL2_86_2A08_G3F1 Direction: sense

TGACATTTTCCAGGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACACACCAGAAGACATCCTCAGGAACAAA
GGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCG
ACTAACTACATTGG

SEQ ID NO:478

>GL2_86_2A09_G3F1 Direction: sense

GTGACATTATCCAGTGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACACACNCAGAAGACATCCTCAGGAAC
AAAGGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGNCAT
CCGCACTAACTACATTGG

SEQ ID NO:479

>GL2_199_3_M13R Direction: anti-sense

CCAATGTAGTTAGTGCGGATGGCAGGGCTGGAACCATTCACCACGTTGTTGTCAAGGGTGAGGAGGACACTGGTA
GAGCTGGAGCAGCCTTTGTTCTGAGGATGTCTTCTGGTGTGGTTCCAAAGACAAACCTCACATTCTGCAGCACC
CCCTGGAAAATGTCA

SEQ ID NO:480

>GL2_86_2A05_G3F1 Direction: sense

TGACATTTTCCAGGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGGAACACACCAGAAGACATCCTCAGGAACAAA
GGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGACAACAACGTGGTGAATGGTTCCAGCCCTGCCATCCG
ACTAACTACATTGG

SEQ ID NO:481

>gi|4507484|ref|NM_003246.1| Homo sapiens thrombospondin 1 (THBS1), mRNA

GGACGCACAGGCATTCCCCGCGCCCCCTCCAGCCCTCGCCCGCCCTCGCCACCGCTCCCGGGCCGCGCGCTC
CGGTACACACAGGATCCCTGCTGGGCACCAACAGCTCCACCATGGGGCTGGCCTGGGGACTAGGCGTCCT
GTTCTGTATGCATGTGTGTGGCACCAACCGCATTCCAGAGTCTGGCGGAGACAACAGCGTGTGTTGACATC
TTTGAACCTACCGGGGGCCGCGCGCAAGGGGTCTGGGCGCCGACTGGTGAAGGGCCCCGACCCTTCCAGCC
CAGCTTTCCGCATCGAGGATGCCAACCTGATCCCCCTGTGCCTGATGACAAGTTCCAAGACCTGGTGGGA
TGCTGTGCGGGCAGAAAAGGGTTTCTCCTTCTGGCATCCCTGAGGCAGATGAAGAAGACCCGGGGCAGC
CTGCTGGCCCTGGAGCGGAAAGACCACTCTGGCCAGGTCTTCAGCGTGGTGTCCAATGGCAAGGCGGGCA
CCCTGGACCTCAGCCTGACCGTCCAAGGAAAGCAGCAGTGGTGTCTGTGGAAGAAGCTCTCCTGGCAAC
CGGCCAGTGGGAAGAGCATCACCTGTTTGTGTCAGGAAGACAGGGCCAGCTGTACATCGACTGTGAAAAG
ATGGAGAATGCTGAGTTGGACGTCCCCATCCAAAGCGTCTTCACCAGAGACCTGGCCAGCATCGCCAGAC
TCCGCATCGCAAAGGGGGGCGTCAATGACAATTTCCAGGGGGTGCTGCAGAATGTGAGGTTTGTCTTTGG
AACCACACCAGAAGACATCCTCAGGAACAAAGGCTGCTCCAGCTCTACCAGTGTCTCCTCACCTTGAC
AACAACGTGGTGAATGGTTCCAGCCCTGCCATCCGCACTAACTACATTGGCCACAAGACAAGGACTTGC
AAGCCATCTGCGGCATCTCCTGTGATGAGCTGTCCAGCATGGTCTTGGAACCTCAGGGGCCTGCGCACCAT
TGTGACCACGCTGCAGGACAGCATCCGCAAAGTGACTGAAGAGAAACAAAGAGTTGGCCAATGAGCTGAGG
CGGCCTCCCCTATGCTATCACAACGGAGTTCAGTACAGAAATAACGAGGAATGGACTGTTGATAGCTGCA
CTGAGTGTCACTGTGAGAACTCAGTTACCATCTGCAAAAAGGTGTCCTGCCCCATCATGCCCTGCTCCAA
TGCCACAGTTTCTGTATGGAGAATGCTGTCTCGCTGTTGGGCCAGCGACTCTGCGGACGATGGCTGGTCT
CCATGGTCCGAGTGGACCTCCTGTTCTACGAGCTGTGGCAATGGAATTCAGCAGCGCGGCGCTCCTGCG
ATAGCCTCAACAACCGATGTGAGGGCTCCTCGGTCCAGACACGGACCTGCCACATTACAGGAGTGTGACAA
AAGATTTAAACAGGATGGTGGCTGGAGCCACTGGTCCCCGTGGTCATCTTGTCTGTGACATGTGGTGTAT
GGTGTGATCACAAGGATCCGGCTCTGCAACTCTCCAGCCCCCAGATGAATGGGAACCCCTGTGAAGGCG
AAGCGCGGGAGACCAAAGCCTGCAAGAAAGACGCTGCCCATCAATGGAGGCTGGGGTCTTGGTCACC
ATGGGACATCTGTTCTGTACCTGTGGAGGAGGGGTACAGAAACGTAAGTCTCTGCAACAACCCCGCA
CCCCAGTTTGGAGGCAAGGACTGCGTTGGTGTATGTAACAGAAAACAGATCTGCAACAAGCAGGACTGTG
CAATTGATGGATGCTGTGTTCCATCCCTGCTTTGCCGGCGTGAAGTGTACTAGCTACCCTGATGGCAGCTG
GAAATGTGGTGTGTTGTTCCCTGGTTACAGTGGAATGGCATCCAGTGCACAGATGTTGATGAGTGCAAA
GAAGTGCCTGATGCCTGCTTCAACCACAATGGAGAGCACCGGTGTGAGAACACGGACCCCGCTACAAC

GCCTGCCCTGCCCCCCACGCTTCACCGGCTCACAGCCCTTCGGCCAGGGTGTGCAACATGCCACGGCCAA
CAAACAGGTGTGCAAGCCCCGTAAACCCCTGCACGGATGGGACCCACGACTGCAACAAGAACGCCAAGTGC
AACTACCTGGGCCACTATAGCGACCCCATGTACCGCTGCGAGTGCAAGCCTGGCTACGCTGGCAATGGCA
TCATCTGCGGGGAGGACACAGACCTGGATGGCTGGCCCAATGAGAACCTGGTGTGCGTGGCCAATGCGAC
5 TTACCACTGCAAAAAGGATAATTGCCCAACCTTCCCAACTCAGGGCAGGAAGACTATGACAAGGATGGA
ATTGGTGTATGCTGTGATGATGACGATGACAATGATAAAATTCCAGATGACAGGGACAACCTGTCCATTCC
ATTACAACCCAGCTCAGTATGACTATGACAGAGATGATGTGGGAGACCGCTGTGACAACCTGTCCCTACAA
CCACAACCCAGATCAGGCAGACACAGACAACAATGGGGAAGGAGACGCCTGTGCTGCAGACATTGATGGA
GACGGTATCTCAATGAACGGGACAACCTGCCAGTACGTCTACAATGTGGACCAGAGAGACACTGATATGG
10 ATGGGGTTGGAGATCAGTGTGACAATTGCCCTTGGAACACAATCCGGATCAGCTGGACTCTGACTCAGA
CCGATTTGGAGATACCTGTGACAACAATCAGGATATTGATGAAGATGGCCACCAGAACAATCTGGACAAC
TGTCCCTATGTGCCCCAATGCCAACCAGGCTGACCATGACAAAGATGGCAAGGGAGATGCCTGTGACCACG
ATGATGACAACGATGGCATTCTGATGACAAGGACAACCTGCAGACTCGTGCCCAATCCCGACCAAGGA
CTCTGACGGCGATGGTCGAGGTGATGCTTGCAAGATGATTTTGACCATGACAGTGTGCCAGACATCGAT
15 GACATCTGTCTTGAGAATGTTGACATCAGTGAGACCGATTTCCGCCGATTCCAGATGATTCTCTGGACC
CCAAAGGGACATCCCAAATGACCCTAACTGGGTGTACGCCATCAGGGTAAAGAACTCGTCCAGACTGT
CAACTGTGATCCTGGACTCGCTGTAGGTTATGATGAGTTTAATGCTGTGGACTTCAGTGGCACCTTCTTC
ATCAACACCGAAAGGGACGATGACTATGCTGGATTTGTCTTTGGCTACCAGTCCAGCAGCCGCTTTTATG
TTGTGATGTGGAAGCAAGTCACCCAGTCTTACTGGGACACCAACCCACGAGGGCTCAGGGATACTCGGG
20 CCTTTCTGTGAAAGTTGTAACTCCACCAAGGCGCTGGCGAGCACCTGCGGAACGCCCTGTGGCACACA
GGAAACACCCCTGGCCAGGTGCGCACCCCTGTGGCATGACCCCTCGTCACATAGGCTGGAAAGATTTCACCG
CCTACAGATGGCGTCTCAGCCACAGGCCAAAGACGGGTTTTCATTAGAGTGGTGTATGAAGGGAAGAA
AATCATGGCTGACTCAGGACCCATCTATGATAAAACCTATGCTGGTGGTAGACTAGGGTTGTTTGTCTTC
TCTCAAGAAATGGTGTCTTCTCTGACCTGAAATACGAATGTAGAGATCCCTAATCATCAAATTGTTGAT
25 TGAAAGACTGATCATAAACCAATGCTGGTATTGCACCTTCTGGAACATATGGGCTTGAGAAAACCCCAAG
ATCACTTCTCCTTGGCTTCTCTTTCTGTGCTTGCATCAGTGTGGACTCCTAGAACGTGCGACCTGCC
TCAAGAAAATGCAGTTTCAAAAACAGACTCATCAGCATTACAGCTCCAATGAATAAGACATCTTCCAAG
CATATAACAATTTGCTTTGGTTTCTTTTGA AAAAAGCATCTACTTGCTTCAGTTGGGAAGGTGCCCATTC
CACTCTGCCTTTGTACAGAGCAGGGTCTATTGTGAGGCCATCTCTGAGCAGTGGACTCAAAGCATTT
30 TCAGGCATGTGAGCAAGGGAGGACTCATAAGATTAGCAATAACAAACACCCCTGACATCTCTCTCAGG
AACACGGGGAGCAGAGGCCAAAGCACTAAGGGAAGGGCGCATACCCGAGACGATTGTATGAAGGGAATAT
GGAGGAATGTTACATGTTTCGGTACTAAGTCATTTTCAGGGGATTGAAAGACTATTGCTGGATTTTATGA
TGCTGACTGGCGTTAGCTGATTAACCCATGTAAATAGGCACCTTAAATAGAAGCAGGAAGGGAGACAAAG
ACTGGCTTCTGGACTTCTCCTGATCCCCACCTTACTCATCACCTTGCAGTGGCCAGAATTAGGGAAT
35 CAGAATCAAACAGTGTAAGGCAGTGTGGCTGCCATTGCCCTGGTCACATTGAAATTGGTGGCTTCATTC
TAGATGTAGCTTGTGAGATGTAGCAGGAAAATAGGAAAACCTACCATCTCAGTGAGCACCAGCTGCCTC
CCAAAGGAGGGGAGCCGTGCTTATATTTTTATGGTTACAATGGCACAAAATTATTTATCAACCTAACTAA
AACATTCTCTTTCTCTTTTCCGTAATTACTAGGTAGTTTCTAATTCTCTCTTTTGAAGTATGATTT
TTTTAAAGTCTTTACGATGTAAATATTTATTTTACTTATTCTGGAAGATCTGGCTGAAGGATTATTC
40 ATGGAACAGGAAGAAGCGTAAAGACTATCCATGTCATCTTTGTTGAGAGTCTTCGTGACTGTGAAGATTGT
AAATACAGATTATTTATTAACCTCTGTTCTGCCTGGAAATTTAGGCTTCATACGGAAAGTGTGAGAGCA
AGTAGTTGACATTTATCAGCAAATCTCTTGCAAGAACAGCACAAAGGAAATCAGTCTAATAAGCTGCTCT
GCCCTTGTGCTCAGAGTGGATGTTATGGGATTCCTTTTCTCTGTTTATCTTTTCAAGTGAATTAG
TTGGTTATCCATTTGCAAATGTTTTAAATTGCAAAGAAAGCCATGAGGTCTTCAATACTGTTTTACCCCA
45 TCCCTTGTGCATATTTCCAGGGAGAAGGAAAGCATATACACTTTTTCTTTTCATTTTTCCAAAAGAGAAA
AAAATGACAAAAGGTGAACTTACATACAAATATTACCTCATTTGTTGTGTGACTGAGTAAAGAATTTTT
GGATCAAGCGGAAAGATTTAAGTGTCTAACAACTTAAAGCTACTGTAGTACCTAAAAAGTCAGTGTG
TACATAGCATAAAAACTCTGCAGAGAAGTATTTCCCAATAAGGAAATAGCATTGAAATGTTAAATACAATT
TCTGAAAGTTATGTTTTTTTTCTATCATCTGGTATACCATTTGCTTTATTTTTATAAATTATTTTCTCATT
50 GCCATTGGAATAGAATATTAGATTGTGTAGATATGCTATTTAAATAATTTATCAGGAAATACTGCCTGT
AGAGTTAGTATTTCTATTTTTATATAATGTTTGCACACTGAATTGAAGAATTGTTGGTTTTTTCTTTTTT
TTGTTTTTTTTTTTTTTTTTTTTTTTTTTTGTCTTTGACCTCCCATTTTACTATTTGCCAATACCTTTTT
CTAGGAATGTGCTTTTTTTTTGTACACATTTTATCCATTTTACATTCTAAAGCAGTGTAAGTTGTATATT
ACTGTTTCTTATGTACAAGGAACAACAATAAATCATATGGAAATTTATATTT

SEQ ID NO:482

>gi|4507485|ref|NP_003237.1| thrombospondin 1 [Homo sapiens]
MGLAWGLGVFLMHVCGTNRIPESGGDNVFDIFELTGAARKGSGRRLVKGPDPSSPAFRIEDANLIPPV
PDDKFQDLVDVAEKGFLLLASLRQMKKTRGTLALLERKDHSGQVFSVVSNGKAGTLDLSLTQVGKQHV
60 VSVEEALLATGQWKSITLFVQEDRAQLYIDCEKMENAEILDVPIQSVFTRDLASIALRLRIAKGGVNDNFQ
VLQNVRFVFGTTPEDILRNKGCSSSTSVLLTLDNNVNGSSPAIRNTYIGHKTKDLQAICGISDELSSM

VLELRGLRTIVTTLQDSIRKVTENKELANELRRPPLCYHNGVQYRNNEEWTVDSCTECHCQNSVTICKK
VSCPIMPSCSNATVPDGECCPRCWPSPSDSADDGWSWPSEWTSCSTSCGNGIQQRGRSCDSLNNRCEGSSVQT
RTCHIQECDKRFKQDGGWSHWPSSCSVTGCGDVITRILCNPSPPQMNGKPCGEARETKACKKDACP
INGGWGPWSPWDICSVTCGGGVQKRSRLCNNPAPQFGGKDCVGDVTENQICNKQDCPIDGCLSNPCFAGV
5 KCTSYPDGSKWKGACPPGYSGNGIQCTDVDECKEVPDACEFNHNGEHRCENTDPGYNCLPCPPRFTGSQPF
GQGVEHATANKQVCKPRNPCTDGTDCNKNACNYLGHYSDDPMYRCECKPGYAGNGIICGEDTDLGWP
ENLVCVANATYHCKKDNCNPLNSGQEDYDKDGI GDACDDDDNDKI PDDRDNCPFHYNPAQYDYDRDDV
GDRCDNCPYNHNPDAQADTDNNGEGDACAADIDGDGILNERDNCQYVYNVDQRDTMDGVGDQCDNCPLEH
10 NPDQLDSDSDRIGDTCDNNQDIDEDGHQNNLDNCPYVNPANQADHDKDGKGDACDHDDNDGIPDDKDNC
RLVPNPDPQKSDSDGGRGDACKDDFDHDSVPDIDDICPENVDISETDFRRFQMIPLDPKGTSONDPNWVVR
HQQKELVQTVNCDPGLAVGYDEFNAVDFSGTFFINTERDDYAGFVFGYQSSSRFYVVMWKQVTQSYWDT
NPTRAQGYSGLSVKVNSTTGPGEHLRNALWHTGNTPGQVRTLWHDPRHIGWKDFTAYRWRLSHRPKTGF
IRVVMYEGKKIMADSGPIYDKTYAGGRLGLFVFSQEMVFFSDDLKYECDP

SEQ ID NO:483

Transgelin 2

>GL3_30H_PCR_G3F1 Direction: sense

CACTGTACCCGAGGGGAGGCCCCAGTAAAGAAGATCAGGCCTCCACCATGGCCTTCAAGCAGATGGAGCAGATC
TCTCAGTTTCCTG

SEQ ID NO:484

>gi|4507356|ref|NM_003564.1| Homo sapiens transgelin 2 (TAGLN2), mRNA
GCCCTTGCTTGTAGTCAGTGCCTGCTCTCCAGCCCGCTTGAACGCTCCCCGAGCCACCGCCACCCATT
GGAATGGCCAAACAGGGGACCTGCATATGGCCTGAGCCGGGAGGTGCAGCAGAAGATTGAGAAACAATATG
25 ATGCAGATCTGGAGCAGATCCTGATCCAGTGGATCACCACCCAGTGCAGAAAGGATGTGGGCCGGCCCCA
GCCTGGACGCGAGAACTTCCAGAACTGGCTCAAGGATGGCACGGTGCTATGTGAGCTCATTAATGCACTG
TACCCCGAGGGGAGGCCCCAGTAAAGAAGATCCAGGCCTCCACCATGGCCTTCAAGCAGATGGAGCAGA
TCTCTCAGTTTCCTGCAAGCAGCTGAGCGCTATGGCATTAAACACCACTGACATCTTCCAAACTGTGGACCT
CTGGGAAGGAAAGAACATGGCCTGTGTGACGCGACGCTGATGAATCTGGGTGGGCTGGCAGTAGCCCCGA
30 GATGATGGGCTCTTCTCTGGGGATCCCAACTGGTTCCCTAAGAAATCCAAGGAGAATCCTCGGAACCTCT
CAGATAACCAGCTGCAAGAGGGGCAAGAACGTGATCGGGTTACAGATGGGCACCAACCGCGGGCGTCTCA
GGCAGGCATGACTGGCTACGGGATGCCACGCCAGATCCTCTGATCCCAACCCAGGCCTTGCCCTGCCCT
CCCACGAATGGTTAATATATATAGTATATATATTTTAGCAGTGACATTCCCAGAGAGCCCCAGAGCTCT
CAAGCTCCTTTCTGTGACGGGTGGGGGTTGAGCCTGTCTGTACCTCTGAGGTGCCTGTCTGGCATCCTC
35 TCCCCCATGCTTACTAATAACATTCCCTTCCCCATAGCCATCAAACTGGACCAACTGGCCTCTTCTTTT
CCCTGGGACCAAAATTTAGGGGCTCAGTCCCTCACCAGCATGCCCTGGCCTATTCTGTCTCTCTCTT
CCCCCTGGCCTGTTCTGTCTCTGAGCTCTGTGTCTCCGTTTATTCCATGGCTGGGAGTCACTGATGCTG
CCTCTGCCTTCTGATGCTGGACTGGCCTTGCTTCTACAAGTATGCTTCTCCACAGCTGTGGCTGCAGGA
ACTTAATTTATAGGGAGGAGCCTGTGGCAGCTGCTGCCCCAGCCACAGCTGCACTGACTGTGCTCACCAC
40 ACATCTGGGGCAGCCTTCCCTGGCAGGGGCCCCCTCGTGGCTTCTCATTTTCCATTCCCTTCACTGTGGCTA
AGGGGTGGGGTGGAGGGGATGGAGAGGGAGGGCTGCCTACCATGGTCTGGGGCTTGAGGAAGATGAGTTT
TTGATTTAAATAAAGAATTTGTCAATTTTG

SEQ ID NO:485

>gi|4507357|ref|NP_003555.1| transgelin 2; SM22-alpha homolog [Homo sapiens]

MANRGPAYGLSREVQOKIEKQYDADLEQILIQWITTQCRKDVGRPQPGRENFQNWLDGTVLCELINALY
PEGQAPVKKIQASTMAFKQMEQISQFLQAAERYGINTTDIFQTVDLWEGKNMACVQRTLMNLGGLAVARD
DGLFSGDPNWFPKSKENPRNFSNQLQEGKNVIGLQMGTNRGASQAGMTGYGMPRQIL

SEQ ID NO:486

Vigilin

>GL3_43A1_1_M13F Direction: anti-sense

CCGGGCAGCTTCGAGTTGGCTCGCTTGCTGTGATGATAATGGTCTCTGAATTGCTATTCTCTGCTGGAAGGGT
CGATTTTGGTGTGTGCTTTCTTACGAATCTTTTAAATGTTTGCGCCTCCTTTCCCAATGATATTCTTGTGAACT
55 GTTTGAAGATCGGAACAGAAATTGAATAGCTATTTTCCACCAGATCTGCCACCATCTTCTGCATGTATTTTGTGC
ATTTTTCACCTCATTTAGGTCCTCTGAGCTGGACAATGTCACTTTTTTGTGCTGGG

SEQ ID NO:487

>GL3_43A1_1_M13R Direction: sense

CCCAGCACAAAAAAGTGACATTGTCCAGCTCAGAGGACCTAAGAATGAGGTGGAAAAATGCACAAAATACATGCA
GAAGATGGTGGCAGATCTGGTGGAAAATAGCTATTCAATTTCTGTTCCGATCTTCAAACAGTTTCACAAGAATAT
5 CATTGGGAAAAGGAGGCGCAAACATTAAAAAGATTTCGTGAAGAAAGCAACACCAAATCGACCTTCCAGCAGAGAA
TAGCAATTCAGAGACCATTATCATCACAGGCAAGCGAGCCAACCTGCGAAGCTGCCCGG

SEQ ID NO:488

>GL3_43A1_2_M13F Direction: sense

CCCAGGACAAAAAAGTGACATTGTCCAGCTCAGAGGACCTAAGAATGAGGTGGAAAAATGCACAAAATACATGCA
GAAGATGGTGGCAGATCGTGGTGGAAAATAGCTATTCAATTTGCTGTTCCGATTCTTCAAACAGTTTCACAAGAA
10 TATCATTGGGAAAAGGAGGCGCAAACGATTCAAGAAAGATTTCGTGAAGAAAGCAACACCAAATCGACCTTCCAGC
AGAGAATAGCAATTTTCAGAGACCATTATCAGTCACAGGCAAGCGAGCCAACCTGCGAAGGGTTGCCCTGGTCCGCT
AGAAACCCAGCACAGGTGGTTAGATGGATACAAGCGGCCGCTCGACTAGTCTGAGGTCTGATACTTCACTGACTA
15 GTCGTTAAGGGCGAATTCAGCACACTGGCGGCCGCTCGACTAGTGGATCCGAGCTC

SEQ ID NO:489

>GL3_43A1_2_M13R Direction: anti-sense

CCGGGCAGCTTCGCAGTTGGCTCGCTTGCCCTGTGATGATAATGGTCTCTGAATTGCTATTCTCTGCTGGAAGGTC
20 GATTTTGGTGTGCTTTCTTCACGAATCTTTTAAATGTTTGCGCCTCCTTTCCCAATGATATTCTTGTGAAACTG
TTTGAAGATCGGAACAGAAATTGAATAGCTATTTTCCACCAGATCTGCCACCATCTTCTGCATGTATTTTGTGCA
TTTTTCCACCTCATTCCTTAGGTCCTCTGAGCTGGACAATGTCACTTTTTTGTGCTGGG

SEQ ID NO:490

>GL3_43A1_4_M13F Direction: sense

CCCAGCACAAAAAAGTGACATTGTCCAGCTCAGAGGACCTAAGAATGAGGTGGAAAAATGCACAAAATACATGCA
GAAGATGGTGGCAGATCTGGTGGAAAATAGCTATTCAATTTCTGTTCCGATCTTCAAACAGTTTCACAAGAATAT
CATTGGGAAAAGGAGGCGCAAACATTAAAAAGATTTCGTGAAGAAAGCAACACCAAATCGACCTTCCAGCAGAGAA
TAGCAATTCAGAGACCATTATCATCACAGGCAAGCGAGCCAACCTGCGAAGTTGCCCGGCCGAGAAACCCAGCAC
30 AGTGGTTAGATAGATAAAGCGGCCGCTCGACTAGTCTGAGGTCTGATACTCACTGACTGTGTAAGGGCGAATTC
CAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGT

SEQ ID NO:491

>GL3_43A1_4_M13R Direction: anti-sense

CCGGGCAGCTTCGCAGTTGGCTCGCTTGCCCTGTGATGATAATGGTCTCTGAATTGCTATTCTCTGCTGGAAGGTC
35 GATTTTGGTGTGCTTTCTTCACGAATCTTTTAAATGTTTGCGCCTCCTTTCCCAATGATATTCTTGTGAAACTG
TTTGAAGATCGGAACAGAAATTGAATAGCTATTTTCCACCAGATCTGCCACCATCTTCTGCATGTATTTTGTGCA
TTTTTCCACCTCATTCCTTAGGTCCTCTGAGCTGGACAATGTCACTTTTTTGTGCTGGG

SEQ ID NO:492

>GL3_43A_PCR_G3F1 Direction: sense

CCCAGCACAAAAAAGTGACATTGTCCAGCTCAGAGGACCTAAGAATGAGGTGGAAAAATGCACAAAATACATGCA
GAAGATGGTGGCAGATCTGGTGGAAAATAGCTATTCAATTTCTGTTCCGATCTTCAAACAGTTTCACAAGAATAT
CATTGGGAAAAGGAGGCGCAAACATTAAAAAGATTTCGTGAAGAAAGCAACACCAAATCGACCTTCCAGCAGAGAA
45 TAGCAAGTTCAGAGACCATTATCATCACAGGCAAGCGAGCCAACCTGCGAAGCTGCCCGG

SEQ ID NO:493

>gi|4885408|ref|NM_005336.1| Homo sapiens high density lipoprotein binding
protein (vigilin) (HDLBP), mRNA

GAATTCGGGGGGCGAGTAAGCCAGCGGCAGGACCAGCGGGCGGGGCCACAACAAAAGCTGGCAGGCTGA
CAGAGGCGGCCTCAGGACGGACCTTCTGGCTACTGACCGTTTTGCTGTGGTTTTCCCGGATTGTGTGTAG
GTGTGAGATCAACCATGAGTTCCGTTGCAGTTTTGACCCAAGAGAGTTTTGCTGAACACCGAAGTGGGCT
GGTTCGCGAACAAATCAAAGTTGCCACTCTAAATTCAGAAGAGGAGAGCGACCCTCCAACCTACAAGGAT
GCCTTCCCTCCACTTCCTGAGAAAGCTGCTTGCCCTGGAAAGTGCCAGGAACCCGCTGGAGCCTGGGGGA
55 ACAAGATCCGACCCATCAAGGCTTCTGTATCACTCAGGTGTTCCATGTACCCCTGGAGGAGAGAAAATA
CAAGGATATGAACAGTTTGGAGAAGGTGAACAAGCAAAAATCTGCCTTGAGATCATGCAGAGAAGTGGT
GCTCACTTGGAGCTGTCTTTGGCCAAAGACCAAGGCCCTCTCCATCATGGTGTGAGGAAAGCTGGATGCTG
TCATGAAGAGCTCGGAAGGACATTGTTGCTAGACTGCAGACTCAGGCCTCAGCAACTGTTGCCATTCCCAA
AGAACACCATCGCTTTGTTATTGGCAAAAATGGAGAGAAACTGCAAGACTTGGAGCTAAAAACTGCAACC
60 AAAATCCAGATCCCACGCCAGATGACCCAGCAATCAGATCAAGATCACTGGCACCAAGAGGGCATCG

AGAAAGCTCGCCATGAAGTCTTACTCATCTCTGCCGAGCAGGACAAACGTGCTGTGGAGAGGCTAGAAGT
AGAAAAGGCATTCCACCCCTTCATCGCTGGGCCGTATAATAGACTGGTTGGCGAGATCATGCAGGAGACA
GGCACGCGCATCAACATCCCCCACCAGCGTGAACCGGACAGAGATTGTCTTCACTGGAGAGAAGGAAC
AGTTGGCTCAGGCTGTGGCTCGCATCAAGAAGATTTATGAGGAGAAGAAAAAGAAGACTACAACCATTGC
5 AGTGGAAAGTGAAGAAATCCCAACACAAGTATGTTCATTGGGCCCAAGGGCAATTCATTGCAGGAGATCCTT
GAGAGAACTGGAGTTTCCGTTGAGATCCACCCCTCAGACAGCATCTCTGAGACTGTAATACTTCGAGGCG
AACCTGAAAAGTTAGGTGAGGCGTTGACTGAAGTCTATGCCAAGGCCAATAGCTTCACCGTCTCCTCTGT
CGCCGCCCTTCTGGCTTACCGTTTTCATCATTGGCAAGAAAGGGCAGAACCCTGGCCAAAATCACTCAG
10 CAGATGCCAAAGGTTACATCGAGTTCACAGAGGGCGAAGACAAGATCACCTTGGAGGGGCCCTACAGAGG
ATGTCAATGTGGCCAGGAACAGATAGAAGGCATGGTCAAAGATTTGATTAACCGGATGGACTATGTGGA
GATCAACATCGACCACAAGTTCCACAGGCACCTCATTGGGAAGAGCGGTGCCAACATAAACAGAATCAAA
GACCAGTACAAGGTGTCCGTGCGCATCCCTCCTGACAGTGAGAAGAGCAATTTGATCCGCATCGAGGGGG
ACCCACAGGGCGTGCAGCAGGCCAAGCAGAGCTGCTGGAGCTTGCACTCTCGCATGGAAAATGAGCGTAC
CAAGGATCTTAATCATTTGAGCAAAGATTTTCATCGCACAATCATTTGGGCAGAAGGGTGAACGGATCCGTGAA
15 ATTCGTGACAAATTTCCAGAGGTTCATCATTAACTTTCCAGACCCAGCACAAAAAGTGACATTGTCCAGC
TCAGAGGACCTAAGAATGAGGTGGA AAAATGCACAAAATACATGCAGAAGATGGTGGCAGATCTGGTGA
AAATAGCTATTCAATTTCTGTTCCGATCTTCAAACAGTTTCAAGAATATCATTGGGAAAGGAGGCGCA
AACATTA AAAAGATTCTGTGAAGAAAGCAACACCAAAATCGACCTTCCAGCAGAGAATAGCAATTCAGAGA
CCATTATCATCACAGGCAAGCGAGCCAACCTGCGAAGCTGCCCCGAGCAGGATTCTGTCTATTTCAGAAAGA
20 CCTGGCCAAACATAGCCGAGGTAGAGGTCTCCATCCCTGCCAAGCTGCACAACTCCCTCATTGGCACCAAG
GGCCGTCTGATCCGCTCCATCATGGAGGAGTGCGGCGGGGTCCACATTCACTTTCCCGTGAAGGTTTCAG
GAAGCGACACCGTTGTTATCAGGGGCCCTTCTCGGATGTGGAGAAGGCCAAGAAGCAGCTCCTGCATCT
GGCGGAGGAGAAGCAAAACCAAGAGTTTCACTGTTGACATCCGCGCCAAGCCAGAATACCACAAATTCCTC
ATCGGCAAGGGGGGCGGCAAAATTCGCAAGGTGCGCGACAGCACTGGAGCACGTGTCTCTTCCCTGCGG
25 CTGAGGACAAGGACCAGGACCTGATCACCATCATTGGAAAGGAGGACGCCGTCCGAGAGGCACAGAAGGA
GCTGGAGGCCCTTGATCCAAAACCTGGATAATGTGGTGAAGACTCCATGCTGGTGGACCCCCAAGCACCAC
CGCCACTTCGTTCATCCGCAGAGGCCAGGTCTTGCGGGAGATTGCTGAAGAGTATGGCGGGGTGATGGTCA
GCTTCCACGCTCTGGCACACAGAGCGCAAAAGTCAACCTCAAGGGCGCCAAGGACTGTGTGGAGGCAGC
CAAGAAACGCATTTCAGGAGATCATTGAGGACCTGGAAGCTCAGGTGACATTAGAATGTGCTATACCCAG
30 AAATTCATCGATCTGTTCATGGGCCCAAAGGTTCCAGAATCCAGCAGATTACTCGGGATTTTCAGTGTTC
AAATTAATTTCCAGACAGAGAGGAGAAAGCAGTTTCACAGTACAGAGCCAGTTGTTCCAGGAGAATGGGGA
CGAAGCTGGGGAGGGGAGAGAGGCTAAAGATTGTGACATCCGCGCTCTCCAAGGAGGTGTGATCATCATC
ATCTCTGGCCGGAAGAAAAGTGTGAGGCTGCCAAGGAAGCTCTGGAGGCATTGGTTCTGTACCATTTG
AAGTAGAGGTGCCCTTTGACCTTACCGTTACGTTATTGGGCAGAAAGGAAGTGGGATCCGCAAGATGAT
35 GGATGAGTTTGAGGTGAACATACATGTCCGGGCACCTGAGCTGCAGTCTGACATCATCGCCATCACGGGC
CTCGTGTCAAATTTGGACCGGGCCAAGGCTGGACTGCTGGAGCGTGTGAAGGAGCTACAGGCCGAGCAGG
AGGACCGGGCTTTAAGGAGTTTTAAGCTGAGTGTCACTGTAGACCCCAAATACCATCCCAAGATTATCGG
GAGAAAGGGGGCAGTAATTACCCAAATCCGTTGGAGCATGACGTGAACATCCAGTTTCTGTATAAGGAC
GATGGGAACCAAGCCCCAGGACCAAAATTACCATCACAGGGTACGAAAAGAACACAGAAGCTGCCAGGGATG
40 CTATCTGAGAATTGTGGGTGAACCTTGAGCAGATGGTTCTGAGGACGTTCCGCTGGACCAAGCGTTCA
CGCCGCATCAATTGTTGTCGCGGCAAGGCCAAGATCATGGACGAATTCAGGTGACCATTCATCGC
TTCCACAGAGCGGAGCCCCAGACCCCAACTGCGTCACTGTGACGGGGCTCCAGAGAATGTGGAGGAAG
CCATCGACCACATCTCAATCTGGAGGAGGAATACCTAGCTGACGTGGTGGACAGTGAGGCGCTGCAGGT
ATACATGAAACCCCCAGCACACGAAGAGGCCAAGGCACCTTCCAGAGGCTTTGTGGTGGCGGACGCACCC
45 TGGACCGCCAGCAGCAGTGAAGGCTCCTGACATGAGCAGCTCTGAGGAATTTCCAGCTTTGGGGCTC
AGGTGGCTCCCAAGACCCTCCCTTGGGGCCCCAAACGATAATGATCAAAAAGAACAGAACCCTCTCCAGC
CTGCTGACCCGAACCAACACACAATGGTTTGTCTCAATCTGACCCAGCGGTGGACCTCCGTAAATT
GTTGAGCGCTCTTCCCTTCCCGAGGTCCGAGGGAGCCTAGCGCTGGCTGTGTGTGCGGCCGCTCCTC
CAGGCCTGGCCGTGCCCGCTCAGGACCTGCTCCACTGTTTAACAATAAACCAAGGTGATGAGCATTGAG
50 CTAAGATAACAGTCCAGCTCCTGGTCCACCCGCATGTGAGTCACTGAGTCACTGAGCAGTCTGGCCTTCATCAGGAG
CTCCGCAGCCGTGGCTAGGATTCCACTTCTGTGTGATGACCTCAGGAAATAAACGTCCTTGACTTTATA
AAAGCCCCGAATTC

SEQ ID NO:494

55 >gi|4885409|ref|NP_005327.1| high density lipoprotein binding protein;
vigilin [Homo sapiens]
MSSVAVLQESFAEHRSLVPQQIKVATLNSSEESDPPTYKDAFPPLPEKAACLES AQEPAGAWGNKIRP
IKASVITQVFHVPLEERKYKDMNQFGEGEQAKICLEIMQRTGAHLELSLAKDQGLSIMVSGKLDVAMKAR
KDIVARLQTQASATVAIPKEHHRFVIGKNGEKLQDLELKTATKIQIPRPPPSNQIKITGTKEGIEKARH
60 EVLLISAEQDKRAVERLEKAFHPFIAGPYNRLVGEIMQETGTRINIPPPSVNRTEIVFTGEKEQLAQ
VARIKKIYEEKKKTTTIAVEVKKSQHKYVIGPKNSLQEI LERTGVSV EIPPSDSIS ETVILRGEPEKL

5 GQALTEVYAKANSFTVSSVAAPSWLHRFIIIGKKGQNLAKITQQMPKVHIEFTEGEDKITLEGPTEDVNVA
QEQIEGMVKDLINRMDYVEINIDHKFHRHLIGKSGANINRIKDQYKVSVRIPPDSEKSNLIRIEGDPQGV
QQAKRELLELASRMENERTKDLIIIEQRFHRTIIIGQKGERIREIRDKFPEVIINFPDPAQKSDIVQLRGP
NEVEKCTKYMQKMAVADLVENSYSISVPIFKQFHKNIIGKGGANIKKIREESNTKIDLPAENSNSSETIIIT
10 GKRANCEAARSRLSIQKDLANIAEVEVSIPAKLHNSLIGTKGRLIRSIMEECGGVHIHFPVEGSGSDTV
VIRGPSSDVEKAKKQLLHLAEKQTKSFTVDIRAKPEYHKFLIGKGGKIRKVRDSTGARVIFPAAEDKD
QDLITTIIGKEDAVREAQKELEALIQLNDNVVEDSMLVDPKHHRHFVIRRGQVLREIAEEYGGVMVSFPRS
GTQSDKVTLLGAKDCVEAAKKRIQEIIEDLEAQVTLECAIPQKFHRSVMGPKGSRIQQITRDFSQIKFP
DREENAVHSTEPVQENGDEAGEGREAKDCDPGSPRRCDIIISGRKEKCEAAKEALEALVPVTIEVEVP
15 FDLHRYVIGQKSGIRKMMDEFEVNIHVPAPELQSDIIATGLAANLDRAKAGLLERVKELQAEQEDRAL
RSFKLSVTVPKYHPKIIIGRKGAUITQIRLEHDVNIQFPDKDDGNQPDQITITGYEKNTEAARDAILRI
VGELEQMVSEDPVLDHRVHARIIGARGKAIRKIMDEFKVDIRFPQSGAPDPNCVTVTGLPENVEAIDHI
LNLEEEYLADVVDSEALQVYMKPPAHEEAKAPSRGFVVRDAPWTASSEKAPDMSSEEFPSFGAQVAPK
TLPWGPKR

SEQ ID NO:495

Vimentin

>G3_8_04_PCR_G3F1 Direction: N/A

20 GCCGTCCAGGCCATCGCCACCCTCCGCAGCCATGTCCACCAGGTCCGTGTCTCTCGTCTCTACCCGAGATGTTT
GGCGGCCCGGCCACCGCGAGCCGGGCGAGCTCCAGCCGGAGCTACGTGACTACGTCCACCCGCACCTACAGCCT
GGCAGCGCGCTGCGCCCGCAGCAGCGCAGCCTCTACGCCTCGTCCCCGGGCGCGGTGTATGCCACGCGCTCC
TCTGCCGTGCGCTGCGGAGCAGCGTGCCCGGGTGCGCTCCTGCGAGACTCGGTGGACTTCTCGCTGGCCGAC
GCCATCAACACCGAGTTCAAGAACACCCGCACCAACAGAAGGTGGAGCTGCAGGAGCTGAATGACCGCTTCGCCA
25 CTATATCGACAGGTGCGCTTCTGAGCAGCAGAATAAGATCCTGCTGGCCGAGCTCGAGCAGCTCAAGGGCCAA
GGCAAGTCGCGCTGGGGACCTCTACGAGGAGGAGATGCGGGAGCTGCGCCGGC

SEQ ID NO:496

>gi|4507894|ref|NM_003380.1| Homo sapiens vimentin (VIM), mRNA

30 GGGCGCGCCAGAGACGCAGCCGCGCTCCCAACACCCACACCCACCGCGCCCTCGTTGCGCTCTTCTCCGG
GAGCCAGTCCGCGCCACCGCCGCGCCAGGCCATCGCCACCTCCGCAGCCATGTCCACCAGGTCCGTG
TCCTCGTCTCTACCGCAGGATGTTTCGGCGGCGCGGCGCCGCGAGCCGCGGAGCTCCAGCCGGAGCT
ACGTGACTACGCTCACCCGCACCTACAGCCTGGGCGAGCGCGCTGCGCCCCAGCACCAGCCGAGCCTCTA
CGCCTCGTCCCCGGGCGGCGGTGTATGCCACGCGCTCCTCTGCGCTGCGCCTGCGGAGCAGCGTCCCGGG
35 GTGCGGCTCCTGCAGGACTCGGTGGACTTCTCGCTGGCCGAGCCATCAACACCGAGTTCAAGAACACCC
GCACCAACGAGAAGGTGGAGCTGCAGGAGCTGAATGACCGCTTCGCCAACTACATCGACAAGGTGCGCTT
CCTGGAGCAGCAGAATAAGATCCTGCTGGCCGAGCTCGAGCAGCTCAAGGGCCAAGGCAAGTCGCGCCTA
GGGGACCTCTACGAGGAGGAGATGCGGGAGCTGCGCCGGCAGGTGGACCAGCTAACCAACGACAAAGCCC
GCGTCGAGGTGGAGCGCGACAACCTGGCCGAGGACATCATGCGCCTCCGGGAGAAATGCGAGGAGGAGAT
GCTTCAGAGAGAGGAAGCCGAAAACACCTGCAATCTTTCAGACAGGATGTTGACAATGCGTCTCTGGCA
40 CGTCTTGACCTTGAACGCAAGTGGAATCTTTGCAAGAAGAGATTGCGCTTTTGAAGAACTCCACGAAG
AGGAAATCCAGGAGCTGCGGCTCAGATTAGGAACAGCATGTCCAAATCGATGTGGATGTTTCCAAGCC
TGACCTCACGGCTGCGCTGCGTGACGTACGTGACCAATGAAAGTGTGGCTGCCAAGAACCCTGCGAGGAG
GCAGAAGAATGGTACAAATCCAAGTTTGCTGACCTCTCTGAGGCTGCCAACCGGAACAATGACGCCCTGC
45 GCCAGGCAAAGCAGGAGTCCACTGAGTACCGGAGACAGGTGCAGTCCCTCACCTGTGAAGTGGATGCCCT
TAAAGGAACCAATGAGTCCCTGGAACGCCAGATGCGTGAAATGGAAGAGAACTTTGCCGTTGAAGCTGCT
AACTACCAAGACACTATTGGCCGCCTGCAGGATGAGATTGAGAATATGAAGGAGGAAATGGCTCGTCACC
TTCTGTAATACCAAGACCTGCTCAATGTTAAGATGGCCCTTGACATTGAGATTGCCACCTACAGGAAGCT
GCTGGAAGGCGAGGAGAGCAGGATTTCTCTGCCTCTTCCAACTTTTCTCTCCCTGAACCTGAGGGAACT
AATCTGGATTCACTCCCTCTGGTTGATACCACTCAAAAAGGACATTCTGATTAAGACGGTTGAAACTA
50 GAGATGGACAGGTTATCAACGAAACTTCTCAGCATCAGATGACCTTGAATAAAAAATGCACACACTCAG
TGGCAGGCGATATATTACCCAGGCAAGAATAAAAAAGAAATCCCATATCTTAAAGAAACAGCTTTCAAGT
GCCTTTCTGCAGTTTTTTCAGGAGCGCAAGATAGATTTGGAATAGGAATAAGCTCTAGTTCTTAACAACCG
ACACTCCTACAAGATTTAGAAAAAAGTTTACAACATAATCTAGTTTACAGAAAAATCTTGTGCTAGAATA
CTTTTTTAAAGGTATTTTGAATACCATTAAACTGCTTTTTTTTTTTCAGCAAGTATCCAACCAACTGG
55 TTCTGCTTCAATAAATCTTTGAAAACTCCA

SEQ ID NO:497

>gi|4507895|ref|NP_003371.1| vimentin [Homo sapiens]

60 MSTRSVSSSSYRRMFGPGTASRPSSSSRSYVTTSTRYSLGSALRPSTSRSLYASSPGGVYATRSSAVRL
RSSVPGVRLQLQDSVDFSLADAINTEFKNTRTNEKVELQELNDRFANYIDKVRFLQONKILLAELEQLKG

ACCAGTACAGGCCCCAGAACGAAGACGAGCTGGAGCTGCGCGAGGGGGACAGGGTGGATGTCATGCAGCAGTGTG
ACGATGGCTGGTTTGTGGGTGTCTCCCGAGGACCCAGAAATTCGGAACGTTCCCTGGAAATTACGTTGCCCCGG
TGTGAGTGGTCTCCATGGCAACTTGGAGCCAGCCAGGATGGGGTGGGGAGCGGTGGCACTCGTGGGAGGGAGAGG
ACCCCCGCCCACATCCTCCTTCCCCAGGACCTGAGCTCCAGCATCTGCAGACGACCCCCGCAGCCTTTCCCTCG
5 GACCCCCCTCGAAGCCCCCTGGACTGATTCCCACCACGATC

SEQ ID NO:508

>GL3_15J_2_M13F Direction: anti-sense

ACCAGTACAGGCCCCAGAACGAAGACGAGCTGGAGCTGCGCGAGGGGGACAGGGTGGATGTCATGCAGCAGTGT
10 GACGATGGCTGGTTTGTGGGTGTCTCCCGAGGACCCAGAAATTCGGAACGTTCCCTGGAAATTACGTTGCCCCG
GTGTGAGTGGTCTCCATGGCAACTTGGAGCCAGCCAGGATGGGGTGGGGAGCGGTGGCACTCGTGGGAGGGAGAG
GACCCCCGCCCACATCCTCCTTCCCCAGGACCTGAGCTCCAGCATCTGCAGACGACCCCCGCAGCCTTTCCCTC
GGACCCTCGAA

15 SEQ ID NO:509

>GL3_15J_2_M13R Direction: sense

CTTGGGGTCTGGTCTCAATAAGGCCTGGAAATGGTCTCCGCTTCAAAGCAGACCTCTGTATTTGGGGAGTGGGGT
GGGGAGGAAATGAAAGGGCTGTGGGAGGAATGCCTGTGAGTCGTGGGTGGGAATCAGTCCAGGGGGCTTCGAGGG
GGGTCCGAGGGAAAGGCTGCGGGGGTCTGTGAGATGCTGGGAGCTCAGGTCTGGGGAAGGAGGATGTGGGCG
20 GGGGTCTCTCCCTCCCACGAGTGCCACCGCTCCCCACCCCATCTGGCTGGCTCCAAGTTGCCATGGAGACCAC
TCACACCGGGGCAACGTAATTTCCAGGGAACGTTCCGAATTTCTGGGTCTCCGGGAGACACCCACAAACCAGCC
ATCGTCACTGCTGCATGACATCCACCTGTCCCCCTCGCGCAGCTCCAGCTCGTCTTCGTTCTGGGGCCTGTA
CTGGT

25 SEQ ID NO:510

>GL3_15J_3_M13F Direction: sense

CTTGGGGTCTGGTCTCAATAAGGCCTGGAAATGGTCTCCGCTTCAAAGCAGACCTCTGTATTTGGGGAGTGGGGT
GGGGAGGAAATGAAAGGGCTGTGGGAGGAATGCCTGTGAGTCGTGGGTGGGAATCAGTCCAGGGGGCTTCGAGGG
GGGTCCGAGGGAAAGGCTGCGGGGGTCTGTGAGATGCTGGGAGCTCAGGTCTGGGGAAGGAGGATGTGGGCG
30 GGGTCTCTCCCTCCCACGAGTGC

SEQ ID NO:511

>GL3_15J_4_M13F Direction: N/A

CTTGGGGTCTGGTCTCAATAAGGCCTGGAAATGGTCTCCGCTTCAAAGCAGACCTCTGTATTTGGGGAGTGGGGT
35 GGGGAGGAAATGAAAGGGCTGTGGGAGGAATGCCTGTGAGTCGTGGGTGGGAATCAGTCCAGGGGGCTTCGAGGG
GGGTCCGAGGGAAAGGCTGCGGGGGTCTGTGAGATGCTGGGAGCTCAGGTCTGGGGAAGGAGGATGTGGGCG
GGGTCTCTCCCTCCCACGAGTGCCACCGTCCCCACCCCATCTGGCTGGCTCCAAGTTGCCATGGAGACCACT
CACACCGGGGCAACGTAATTTCCAGGGAACGTTCCGAATTTCTGGGTCTCCGGGAGACACCCACAAACCAGCCA
TCGTACAC

40 SEQ ID NO:512

>GL3_15J_PCR_G3F1 Direction: N/A

CTTGGGGTCTGGTCTCAATAAGGCCTGGAAATGGTCTCCGCTTCAAAGCAGACCTCTGTATTTGGGGAGTGGGGT
GGGGAGGAAATGAAAGGGCTGTGGGAGGAATGCCTGTGAGTCGTGGGTGGGAATCAGTCCAGGGGGCTTCGAGGG
45 GGGTCCGAGGGAAAGGCTGCGGGGGTCTGTGAGATGCTGGGAGCTCAGTCTGGGGAAGGAGGATGTGGGCGG
GGGTCTCTCCCTCCCACAGTGCCACGTCCCCCACCCATCCTGGGCGTGGCTCCAAGTTGCCATGGGAAGAAC
ACTCACCACCGGGGCCAACCGTAATTTCCAAGGGAACCGGTTCCCGAAATTTCTGGGGTCCCTCCGGGAAAA
AACCAACAAAACAAGCCCTACGGCA

50 SEQ ID NO:513

>GL3_37B_PCR_G3F1 Direction: sense

CTTGGGGTCTGGTCTCAATAAGGCCTGGAAATGGTCTCCGCTTCAAAGCAGACCTCTGTATTTGGGGAGTGGGGT
GGGGAGGAAATGAAAGGGCTGTGGGAGGAATGCCTGTGAGTCGTGGGTGGGAATCAGTCCAGGGGGCTTCGAGGG
GGGTCCGAGGGAAAGGCTGCGGGGGTCTGTGAGATGCTGGGAGCTCAGTCTGGGGAAGGAGGATGTGGGCGG
55 GGGTCTCTCCCTCCCACGAGTGCCACCGTCCCCACCCCATCCTGGCTGGCTCCAAGTTGCCATGGAGACCACT
CACACCGGGGCAACGTAATTTCCAGGGAACGTTCCGAATTTCTGGGTCTCCGGGAGACACCCACAAACCAGCCA
TCGTCACTGCTGCATGACATCCACCTGTCCCCCTCGCGCAGCTCCGCTCGTCTTCGTTCTGGGGCCTGTACT
GGT

SEQ ID NO:514

>GL3_37D_1_M13R Direction: sense

CTTGGGGTCTGGTCTCAATAAGGCCTGGAAATGGTCTCCGCTTCAAAGCAGACCTCTGTATTTGGGGAGTGGGGT
GGGGAGGAAATGAAAGGGCTGTGGGAGGAATGCCTGTGAGTCGTGGGTGGGAATCAGTCCAGGGGGCTTCGAGGG
GGGTCCGAGGGAAAGGCTGCGGGGGTCTGTCTGCAGATGCTGGGAGCTCAGGTCTTGGGGAAGGAGGATGTGGGCG
GGGGTCTCTCCCTCCCACGAGTGCCACCGTCCCCACCCATCCTGGCTGGCTCCAAGTTGCCATGGAGACCACTC
ACACCGGGCAACGTAATTTCCAGGGAACGTTCCGAATTTCTGGGTCTCCGGGAGACACCCACAAACCAGCCAT
CGTCACACTGCTGCATGACATCCACCCTGTCCCCCTCGCGCAGCTCCAGCTCGTCTTCGTTCTGGGGCCTGTACT
GGTCCGAGAAAACCCAGCACAGTGGTTAGATACGATAAAGCGGACCGCTCGACTAGTCTGAGGTCTGATACTCA
ACTGGACTGTCTGAATGGGGCGAATTTCTTGGCAAGATTTTCCATCACACTTGCGGGGCGTCAAGCATGCAT

SEQ ID NO:515

>GL3_37D_2_M13R Direction: anti-sense

ACCAGTACAGGCCCCAGAACGAAGACGAGCTGGAGCTGCGCGAGGGGGACAGGGTGGATGTCATGCAGCAGTGTG
ACGATGGCTGGTTTTGTGGGTGTCTCCCGGAGGACCCAGAAATTCGGAACGTTCCCTGGAAATTACGTTGCCCCGG
TGTGAGTGGTCTCCATGGCAACTTGGAGCCAGCCAGGATGGGGTGGGGAGCGGTGGCACTCGTGGGAGGGAGAGG
ACCCCCGCCACATTCTCCTTTCCCAGGACCTGAGCTCCAGCATCTGCAGACGAACCCCGGAAGCTTTCCCTCG
GA

SEQ ID NO:516

>GL3_37D_3_M13R Direction: N/A

CTTGGGGTCTGGTCTCAATAAGGCCTGGAAATGGTCTCCGCTTTCAAAGCAGACCTCTGTATTTGGGGAGTGGG
GTGGAGGAGGAAATGAAAGGGCTGTGGGAGGAATGCCTGTGAGATCGGTGGGTGGGAATCAGTCCAGGGGGCTT
CGAGGGGGTCCGAGGGAAAGGCTGCGGGGGTCTGTCTGCAGATCCTGGGAGCCTCAGGTCTTGGGGAAGGAGGAT
GTGGGCGGGGGTCTCTCCCTCCCACGAGTGCCACCGTCCCCACCCATCCTGGCTGGCTCCAAGTTGCCATGG
AGACCACTCACACCGGGGCAACGTAATTTCCAGGGAACGTTCCGAATTTCTGGGTCTCCGGGAGAACACCCACA
AACCAGCCATCGTCACACTGCTGCATGACATCCACCCTGTCCCCCTCGCGCAGCTCCCAGCTCGTCTCCGTNCTG
GGCCTGTACTGGTC

SEQ ID NO:517

>GL3_37D_4_M13R Direction: sense

CTTGGGGTCTGGTCTCAATAAGGCCTGGAAATGGTCTCCGCTTCAAAGCAGACCTCTGTATTTGGGGAGTGGGGT
GGGGAGGAAATGAAAGGGCTGTGGGAGGAATGCCTGTGAGTCGTGGGTGGGAATCAGTCCAGGGGGCTTCGAGGG
GGGTCCGAGGGAAAGGCTGCGGGGGTCTGTCTGCAGATGCTTGGGAGCTCAGGTCTTGGGGAAGGGGGATGTGGGC
GGGGTCTCTCCCTCCCACGAGTGCCACCGTCCCCAACCCATCCTGGCTGGCTCCAAGTTGCCATGGAGACCAC
TCACACCGGGCAACGTAATTTCCAGGGAACGTTCCGAATTTCTGGGTCTCCGGGAGACACCCACAAACCAGCC
ATCGTCGCACTGCTGCATGACATCCACCCTGTCCCCCTCGCGCAGTTCAGCTCGTCTTACGTTCTGGGGCCTGT
AACTGGGT

SEQ ID NO:518

>GL3_37D_PCR_G3F1 Direction: sense

GCTTGGGGTCTGGTCTCAATAAGGCCTGGAAATGGTCTCCGCTTCAAAGCAGACCTCTGTATTTGGGGAGTGGGG
TGGGGAGGAAATGAAAGGGCTGTGGGAGGAATGCCTGTGAGTCGTGGGTGGGAATCAGTCCAGGGGGCTTCGAGG
GGGTCCGAGGGAAAGGCTGCGGGGGTCTGTCTGCAGATGCTGGGAGCTCAGTCTTGGGGAAGGAGGATGTGGGCG
GGGGTCTCTCCCTCCCACGAGTGCCACCGTCCCCACCCATCCTGGCTGGCTCCAGTTGCCATGAACCACTCA
CACCGGGGCAACGTAATTTCCAGGGAACGTTCCGAATTTCTGGGTCTCCGGGAGACACCCACAAACCAGCCATC
GTCACACTGCTGCATGACATCCACCCTGTCCCCCTCGCGCAGTTCAGCTCGTCTTACGTTCTGGGGCCTGTACTGG
TCCGCAACCCAGCACAGTGGTTAGATGATTAAGCGGCCGCTCGACTAGTCTGAGGTCTGATACTCACTGACTTT
TTAA

SEQ ID NO:519

>gi|19923334|ref|NM_005775.2| Homo sapiens vinexin beta (SH3-containing adaptor molecule-1) (SCAM-1), mRNA

CGGACGCGTGGGCAGGCAGCAGCCGGGCAGGGATGCTCCTGCGCTCCCGGGCGGCCTCGGGCCAGCCAC
CTGCTCGCCGGGGAAGAGGACACGCAGAGGAGCAGCTGGCTTGCCCGGAGTCTCTCCACCTTGACCCAAG
CATGCAGGGCCACCCCGCAGCCTCCGCGCTGGGCTCAGCCTGGACGACTTCATCCCTGGCCACCTCCAG
TCCCACATAGGCTCTTCTCCCGGGGACACGGGTGCCCGTGATCCGGAATGGTGGCTCCAACACCTTA
ATTTCCAGTTCCACGACCCCGCGCCAGGACTGTGTGCAATGGGGGCTACACACCAAGACGAGATGCTTC
CCAGCACCCGGACCTGCGTGGTATCAGACCTGGCCAGGCCCTGGGAGCAAGCCCTCTGCAAGCACAAAG
ATCCCTGCCTCCAGCACACCCAGAACTGGTCAGCCACGTGGACCAAGGACAGCAAGCGTCGGGACAAGC

GCTGGGTCAAGTACGAGGGAATCGGGCCCGTGGACGAGAGCGGCATGCCCATTTGCCCCCGATCCAGCGT
 TGACAGACCCAGAGACTGGTACCGGAGAATGTTCCAGCAGATTACCGGAAAATGCCAGACTTGACAGCTG
 GACTGGACCTTCGAGGAGCCACCCAGAGACCCAGGCATCTAGGAGGCCAGCAAAGACCTGCCACAGGC
 CCGGCCCGGCAACATCTTCCAGTGGAAAGAGCTGGGACCACTCTGAAGAGTTACCTAGAAGCACCTTCAA
 5 CTACAGACCTGGAGCATTCTCCACTGTGCTGCAGCCCTCAAATCAGGTGCTCAGACGCCGGGAAAAAGTA
 GACAATGTCTGGACGGAAGAGTCTTGAACCAAGTTTCTGCAGGAAGTAGAGACTGGGAGAGAGAGCTGGC
 AACCGCTGGTGGACGACCTGGTGAGAAGCCCTCCAGCCCATTGAGGTGCTGCTGGAGAGAGAGCTGGC
 CGAGCTGAGCGCCGAGCTGGACAAGGACCTGCGGGCAATTGAGACCCGACTGCCGTCCCCCAAGAGCTCG
 10 CCGGCGCCCCGACGGGCCCCGGAGCAGCGGCCCCGGCCGGCCCGCCTCAGCCTGGAGCTCCAGCTACC
 CACATGCACCTTACCTGGGTTCCGCCCCGGTCCCTGAGTCCCCACAAAATGGCTGATGGAGGAAGCCCCCTT
 CCTAGGTGGAGGAGACTTTGTCTACCTTCTCAACCCGAGACCCCTAGTGCCTCTAACGGAGGGGGCAGC
 CCAGCCAGGAGGAAGAGAAGAAGAGAAAGGCCCGCAGGCTCAAGTTTGACTTCCAGGCGCAGTCCCCCA
 AGGAGCTGACTCTGCAGAAGGGTGACATTGTCTACATCCACAAGGAGGTGGACAAGAAGTGGCTGGAGGG
 AGAGCACCACGGCCGCCTGGGCATCTTCCCTGTCTAATTATGTGGAGGTGCTGCCCGCAGATGAGATCCCT
 15 AAGCCCATCAAGCCCCGACCTACCAGGTGCTGGAGTATGGAGAGGCTGTGGCCAGTACACCTTCAAGG
 GGGACCTGGAGGTGGAGCTGTCTTCCGCAAGGGAGAGCACATCTGCCTGATCCGCAAGGTGAACGAGAA
 CTGGTACGAGGGACGCATCACGGGCACGGGGCGCCAAGGCATATTCCTGCCAGCTACGTGCAGGTGTCT
 CGTGAACCCCGGCTCCGGCTCTGTGACGACGGCCCCCAGCTCCCCACGTCTCCCCGCCTGACCGCTGCCG
 CCGCTCAGCCCGTCAACCCAGCTCCCCCTCAGCCCTGCGCAGCCAGCTGACCCACCGACTTGGGGGG
 20 ACAGACCTCCCCCGCTCGCACTGGCTTCTCCTTCCCCACCCAGGAGCCTAGACCCAGACCCAGAATCTT
 GGCACCCCTGGTCCAGCTCTGTCCCACCTCGAGGTCCCAGCCATCCCCTGGACCTGGGGACCTCCTCTC
 CTAACACCTCTCAGATACACTGGACCTGACCCGCTACCGGGCGATGACCAAGTACAGGCCCCAGAAGACGA
 GCTGGAGCTGCGCGAGGGGGACAGGGTGGATGTCTATGACGAGTGTGACGATGGCTGGTTTGTGGGTGTC
 TCCCGGAGGACCCAGAAATTCGGAACGTTCCCTGGAAATTACGTTGCCCGGTGTGAGTGGTCTCCATGG
 25 CAACTTGGAGCCAGCCAGGATGGGGTGGGGAGCGGTGGCACTCGTGGGAGGGAGAGGACCCCCGCCACA
 TCCTCCTTCCCAGGACCTGAGCTCCCAGCATCTGCAGACGACCCCGCAGCATTTCCCTCGGACCCCCC
 TCGAAGCCCCCTGGACTGATTCCCACCCACGACTCACAGGCATTCCTCCACAGCCCTTTCATTTCTCC
 CCACCCCACTCCCCAAATACAGAGGTCTGCTTTGAAGCGGAGACCATTTCAGGCCTTATTGAGACCAGA
 CCCCAGTCCCCCACCCCCATCCTGCTCCAGCGTTTCTCTAACAGGGACCAGCTCTCCGCTTTGCCCCC
 30 ACGGGGTTCTCTAACAGAACAGCTTCTAGCCTCGTAGAGACCAAAGGCCGCCCGCCGCTGCTGGGG
 TTCCTCCCAGCACCCAGCTTGGCTGGCTGCCCTCTTTGCCCTTCTGGCCTCCAGCTGGGTGTGGGGGGCG
 GAGCAAGGCGGGGACAGCAGCAGCACTTCTTAGCGALETAGGCCTGGCAAGAGCTCTGGCCCCAAGGC
 CTCCTCTTCCCAGGGGCTGCCAAGTCTGGCCCTGGCCCTGGCATATCACCCCGCACTGTGGGGGCCAGGC
 ACCACTAGCCTGGCTCAAATATTCCCAGGGAGACTGCTGTGTGCTGCCCGCCTGCCTGCTGGCTCTCC
 35 CCAGCCCCACATCCCCTCTGGAAGAGAATGTAAATAAACCTGGACACAAGGGCCTTCGTGGCCTCGA

SEQ ID NO:520

>gi|19923335|ref|NP_005766.2| vinexin beta (SH3-containing adaptor
 molecule-1) [Homo sapiens]
 40 MQGPPRSLRAGLSLDDFIPGHLQSHIGSSSRGTRVPVIRNGGSNTLNFQFHDPAVRTVCNGGYTPRRDAS
 QHPDPAWYQTPGPGSKPSASTKIPASQHTQNSATWTKDSKRDRKRWVKYEGIGPVDESGMPIAPRSSV
 DRPRDWYRMFQQIHRKMPDLQLDWTFRMEPRDPRHLGAQQRPAHRPGPATSSSGRSWDHSEELPRSTFN
 YRPGAFSTVLQPSNQVLRREKVDNVWTEESWNQFLQELTGTGQRPKKPLVDDPGEKPSQPIEVLLERELA
 45 ELSAELDOKDLRAIETRLPSPKSSPAPRRAPQRPPAGPASAWSSSYPHAPYLGSARSLSPHKMADGGSPF
 LGRRDFVYPSSTRDPSASNGGSPARREEKKRKAARLKFDFOAQSPKELTLQKGDIVYIHKEVDKNWLEG
 EHHGRLGIFPANYVEVLPADIEPKPIKPPTYQVLEYGEAVAQYTFKGDLEVELSFRKGEHICLIRKVNEN
 WYEGRITGTGRQGIFFASYVQVSREPRLRLCDDGPQLPTSPRLTAAARSARHPSSPSALRSPADPTDLGG
 QTSRRTGFSFPTQEP RPQTQNLGTPGPALSHSRGPSHPLDLGTSSPNTSQIHWTPIRYAMYQYRPQNEDE
 50 LELREGDRDVMQQCDDGWFVGVSRRTQKFGTFPGNYVAVP

SEQ ID NO: 521

NM_004265 (Homo sapiens fatty acid desaturase 2 (FADS2))NM_004265.2
 GI:14141180
 55 1 agggggcgcg gtgggaggag taggagaaga caaaagccga aagcgaagag ggccccgggct
 61 gcacacaccg gctgggaggc agcgtctgt gcagcgagca gccggcgcg ggaggccgca
 121 gtgcacgggg cgtcacagtc ggcagcgagc atgggggaagg gagggaaacca gggcgagggg
 181 gccgcccagc gcgaggtgtc ggtgcccacc ttcagctggg aggagattca gaagcataac
 241 ctgcgcaccg acaggtggct ggtcattgac cgcaagggtt acaacatcac caaatggtcc
 301 atccagcacc cggggggcca gcgggctatc gggcactacg ctggagaaga tgcaacggat
 60 361 gccttccgcg ccttccaccc tgacctgaa ttcgtgggca agttcttgaa acccctgctg
 421 attggtgaac tggccccgga ggagcccagc caggaccacg gcaagaactc aaagatcact

481 gaggacttcc gggccctgag gaagacggct gaggacatga acctgttcaa gaccaaccac
 541 gtgtttcttc tctctctcct ggccacatc atcgccctgg agagcattgc atggttcact
 601 gtcttttact ttggcaatgg ctggattcct accctcatca cggcctttgt ccttgctacc
 661 tctcaggccc aagctggatg gctgcaacat gattatggcc acctgtctgt ctacagaaaa
 5 721 cccaagtgga accaccttgt ccacaaattc gtcattggcc acttaaaggg tgcctctgcc
 781 aactggtgga atcatcgcca cttccagcac caccgccaagc ctaacatctt ccacaaggat
 841 cccgatgtga acatgctgca cgtgtttgtt ctgggcgaat ggcagcccat cgagtacggc
 901 aagaagaagc tgaaataacct gccctacaat caccagcacg aatacttctt cctgattggg
 961 ccgccgctgc tcatcccat gtatttccag taccagatca tcatgacct catcacctac
 10 1021 aagaactggg tggacctggc ctgggcctgc agctactaca tccggttctt catcacctac
 1081 atccctttct acggcatcct gggagccctc cttttcctca acttcatcag gttcctggag
 1141 agccactggg ttgtgtgggt cacacagatg aatcacatcg tcatggagat tgaccaggag
 1201 gcctaccgtg actggttcag tagccagctg acagccacct gcaacgtgga gcagtccttc
 1261 ttcaacgact ggttcagtgg acaccttaac ttccagattg agcaccacct cttccccacc
 15 1321 atgccccggc acaacttaca caagatcgcc ccgctggtga agtctctatg tgccaagcat
 1381 ggcattgaat accaggagaa gccgctactg agggccctgc tggacatcat caggtccctg
 1441 aagaagtctg ggaagctgtg gctggacgcc taccttcaca aatgaagcca cagcccccg
 1501 gacaccgtgg ggaaggggtg caggtggggt gatggccaga ggaatgatgg gcttttgttc
 1561 tgaggggtgt ccgagaggct ggtgtatgca ctgctcacgg accccatgtt ggaatcttct
 20 1621 ccctttcttc tctctttttt cttctcacat ctccccata gcacctgcc ctcattgggac
 1681 ctgccctccc tcagccgtca gccatcagcc atggccctcc cagtgcctcc tagcccttc
 1741 ttccaaggag cagagagggt gccaccgggg gtggctctgt cctacctcca ctctctgccc
 1801 ctaaagatgg gaggagacca gcggtccatg ggtctggcct gtgagtctcc ccttgcagcc
 1861 tggctactag gcatcacccc cgctttgggt cttcagatgc tcttgggggt cataggggca
 25 1921 ggtcctagtc gggcagggcc cctgaccctc ccggcctggc ttcactctcc ctgacggctg
 1981 ccattggtcc accctttcat agagaggcct gctttgttac aaagctcggg tctccctcct
 2041 gcagctcggg taagtaccgg aggcctctct taagatgtcc agggccccag gccgcggggc
 2101 acagccagcc caaaccttgg gccctggaag agtccctccac cccatcacta gagtgtctct
 2161 accctgggct ttcacgggcc ccattccacc gctccccaa cttgagcctg tgaccttggg
 30 2221 accaaagggg gagtccctcg tctcttgtga ctcagcagag gcagtggcca cgttcaggga
 2281 ggggcgggct ggcctggagg ctcagccccc cctccagctt ttcctcaggg tgtcctgagg
 2341 tccaagattc tggagcaatc tgacccttct ccaaaggctc tgttatcagc tgggcagtgc
 2401 cagccaatcc ctggccattt ggccccaggg gacgtggggc ctgcaggctg caggagggca
 2461 ctggagctgg gaggtctcgt cccagccctc cccatctcgg ggctgctgtg tggacggcgc
 35 2521 tgcctcaggc actctcctgt ctgaacctgc ccttactgtg tttaacctgt tgetccagga
 2581 tgcattctga taggaggggg cggcagggct gggccttgtg acaatctgcc tttcaccaca
 2641 tggccttgcc tgggtggccc tgactgtcag ggaggggcag ggaggcagag cgggagggag
 2701 tctcaggagg aggtcgccct gaggggctgg ggagggggta cctcatgagg accagggtgg
 2761 agctgagaag agggagggtg gggggctgga ggtgctggta gctgagggga cgggcaagt
 40 2821 agagggagg gagggaaagt ctgggaggat cctgagctgc tgttgcagtc taaccacta
 2881 atcagttctt agattcaggg gaagggcagg caccaacaac tcagaatggg ggctttcggg
 2941 gagggcgcc agtcccccca gctctaagca gccaggaggg acctgcatct aagcatctgg
 3001 gttgccatgg caatggcatg cccccagct actgtatgcc cccgaccccc gcagaggcag
 3061 aatgaacca tagggagctg atcgtaatgt ttatcatgtt acttccccac cctacattt
 45 3121 tttgaaataa aataaggaat tttattctc

SEQ ID NO:522

NP_004256 (fatty acid desaturase 2; linoleoyl-CoA desaturase (delta-6-
 desaturase)-like 2; delta-6 fatty acid desaturase) [Homo
 sapiens].NP_004256.1 GI:4758334

50 1 mgkgnqgeg aaerevsvpt fsweeiqkhn lrtdrwlvid rkvnitkws iqhpggqrvi
 61 ghyagedatd afrahpdl fvgkflkpl igelapeeps qdhgknskit edfralrkta
 121 edmnlfktnh vffllllahi ialesiawft vfyfngwip tlitaflvat sqagagwlqh
 181 dyghlsvyrk pkwnhlvhkf vighlkgasa nwnhrhfqh hakpnifhkd pdvnmhvf
 55 241 lgewqpieyg kkkllkylpyn hqheyfflig ppllipyf yqiimtmivh knwvdlawav
 301 syirffity ipfygilgal lflnfirfle shwfvwtqm nhivmeidqe ayrdwfssql
 361 tatenveqsf fndwfsghln fqiehhlfpt mprhnlhkia plvkslcakh gieyqekpl
 421 ralldiirsl kksghklwlda ylhk

SEQ ID NO:523

AF134404 (Homo sapiens delta-6 fatty acid desaturase (CYB5RP) mRNA),
AF134404.1 GI:4868365

```

5      1 cttcgcttcc ctcgggggtct tgctcggacc tcggccaccg cctgggatcc ccaggactcg
      61 tgcgtgcagc atgggcggcg tcggggagcg gggaccgcgg gagggaccgc cgcagccggg
      121 ggcaccgctg cccaccttct gctgggagca gatccgcgcg caccgaccag cccgcgacaa
      181 gtggctggtc atcgagcgcc gcgtctacga catcagccgc tgggcacagc ggcacccagg
      241 gggcagccgc ctcatcggcc accacggcgc tgaggacgcc acggatgcct tccgtgcctt
      301 ccatcaagat ctcaattttg tgcgcaagtt cctacagccc ctggttgatt gagagctggc
10     361 tccggaagaa cccagccagg atggaccctt gaatgcgcag ctggtcgagg acttccgagc
      421 cctgcaccag gcagccgagg acatgaagct gtttgatgcc agtcccacct tctttgcttt
      481 cctactgggc cacatcctgg ccatggaggt gctggcctgg ctcttatctt acctcctggg
      541 tcctggctgg gtgcccagtg ccctggccgc cttcatcctg gccatctctc aggetcagtc
      601 ctggtgtctg cagcatgacc tgggcatgc ctccatcttc aagaagtcct ggtggaacca
15     661 cgtggcccag aagtctgtga tggggcagct aaagggttcc tccgccactt ggtggaactt
      721 ccgccacttc cagcaccacg ccaagcccaa catcttccac aaagaccag acgtgacggg
      781 ggcgcccgtc ttctcctctg gggagtcata cgctcagtat ggcaagaaga aacgcagata
      841 cctaccctac aaccagcagc acctgtactt cttcctgatc ggcccgcgcg tgcacacctt
      901 ggtgaacttt gaagtggaaa atctggcgta catgctgggtg tgcatgcagt gggcggattt
20     961 gctctggggc gccagcttct atgcccgctt cttcttatcc tacctccctt tctacggcgt
      1021 ccctgggggtg ctgctcttct ttggtgctgt cagggtcctg gaaagccact ggttcgtgtg
      1081 gatcacacag atgaaccaca tccccaaagg gatcggccac gagaagcacc gggactgggt
      1141 cagctctcag ctggcagcca cctgcaacgt ggagccctca cttttcacca actggttcag
      1201 cgggcacctc aacttcagaa tcagacacca cctcttcccc aggatgccga gacacaacta
25     1261 cagccgggtg gccccgctgg tcaagtcgct gtgtgccaa gacggcctca gctacgaagt
      1321 gaagcccttc ctcaccgcgc tgggtggacat cgtcaggtcc ctgaagaagt ctggtgacat
      1381 ctggctggac gcctacctcc atcagtgaag gcaacaccca ggccggcaga gaagggtcga
      1441 gggcaccagc aaccaagcca gccccggcg ggatcgatac cccacccctt cactggcca
      1501 gcctgggggt gcactgcctg ccctcctggg actgttgtct tcccctcggc cccctcacat
30     1561 gtgtattcag cagccctatg gccttggtc tgggcctgat gggacagggg tagagggag
      1621 gtgagcatag cacattttcc tagagcgaga attgggggaa agctgttatt tttatattaa
      1681 aatacattca gatgtaaaaa aaaaa

```

SEQ ID NO:524

35 AAD31282 (delta-6 fatty acid desaturase) [Homo sapiens]. AAD31282.1
GI:4868366

```

      1 mggvgppgr egpaqpgapl ptfcweqira hdqpgdkwlv ierrvydisr waqrhpggsr
      61 lighhgaeda tdafrafhqd lnfvrkflqp lligelapee psqdgplnaq lvedfralhq
      121 aaedmklfda sptffafllg hilamevlaw lliyllgpgw vpsalaafil aisqaqswcl
40     181 qhdlghasif kkswnnhvaq kfvmgqlkgf sahwnfrhf qhhakpnifh kdpdvtvapv
      241 flgessvey gkkkrlylpy nqghlyffli gpplltlvnf evenlaymlv cmqwadllwa
      301 asfyarffls ylpfygvpv llffvavrvi eshwfvwitq mnhipkeigh ekhrdwvssq
      361 laatcnveps lftnwfsghl nfqiehhlf rmp rhnysrv aplvkslcak hglseyekpf
45     421 litaldivrs lkksqdiwld aylhq

```


**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)